

A DISCONTINUOUS GALERKIN TIME-STEPPING SCHEME FOR THE VELOCITY TRACKING PROBLEM*

EDUARDO CASAS[†] AND KONSTANTINOS CHRYSAFINOS[‡]

Abstract. The velocity tracking problem for the evolutionary Navier–Stokes equations in two dimensions is studied. The controls are of distributed type and are submitted to bound constraints. First and second order necessary and sufficient conditions are proved. A fully discrete scheme based on the discontinuous (in time) Galerkin approach, combined with conforming finite element subspaces in space, is proposed and analyzed. Provided that the time and space discretization parameters, τ and h , respectively, satisfy $\tau \leq Ch^2$, then L^2 error estimates of order $O(h)$ are proved for the difference between the locally optimal controls and their discrete approximations.

Key words. evolution Navier–Stokes equations, optimal control, a priori error estimates, discontinuous Galerkin methods

AMS subject classifications. 49J20, 49K20, 35K55, 65N30

DOI. 10.1137/110829404

1. Introduction. In this paper we prove some error estimates for the numerical approximation of a distributed optimal control problem governed by the evolution Navier–Stokes equations with pointwise control constraints. More precisely, we consider the following problem:

$$(P) \quad \begin{cases} \min J(\mathbf{u}) \\ \mathbf{u} \in \mathcal{U}_{ad}, \end{cases}$$

where

$$\begin{aligned} J(\mathbf{u}) = & \frac{1}{2} \int_0^T \int_{\Omega} |\mathbf{y}_{\mathbf{u}}(t, x) - \mathbf{y}_d(t, x)|^2 dx dt + \frac{\gamma}{2} \int_{\Omega} |\mathbf{y}_{\mathbf{u}}(T, x) - \mathbf{y}_{\Omega}(x)|^2 dx \\ & + \frac{\lambda}{2} \int_0^T \int_{\Omega} |\mathbf{u}(t, x)|^2 dx dt. \end{aligned}$$

Here $\mathbf{y}_{\mathbf{u}}$ denotes the solution of the two-dimensional evolution Navier–Stokes equations

$$(1.1) \quad \begin{cases} \mathbf{y}_t - \nu \Delta \mathbf{y} + (\mathbf{y} \cdot \nabla) \mathbf{y} + \nabla p = \mathbf{f} + \mathbf{u} & \text{in } \Omega_T = (0, T) \times \Omega, \\ \operatorname{div} \mathbf{y} = 0 & \text{in } \Omega_T, \mathbf{y}(0) = \mathbf{y}_0 & \text{in } \Omega, \mathbf{y} = 0 & \text{on } \Sigma_T = (0, T) \times \Gamma, \end{cases}$$

and \mathcal{U}_{ad} is the set of feasible controls, defined for $-\infty \leq \alpha_j < \beta_j \leq +\infty$, $j = 1, 2$, by

$$\mathcal{U}_{ad} = \{\mathbf{u} \in L^2(0, T; \mathbf{L}^2(\Omega)) : \alpha_j \leq u_j(t, x) \leq \beta_j \text{ a.e. } (t, x) \in \Omega_T, j = 1, 2\}.$$

*Received by the editors April 4, 2011; accepted for publication (in revised form) June 14, 2012; published electronically September 18, 2012.

<http://www.siam.org/journals/sinum/50-5/82940.html>

[†]Departamento de Matemática Aplicada y Ciencias de la Computación, E.T.S.I. Industriales y de Telecomunicación, Universidad de Cantabria, Av. Los Castros s/n, 39005 Santander, Spain (eduardo.casas@unican.es). This author's work was partially supported by the Spanish Ministerio de Economía y Competitividad under project MTM2011-22711.

[‡]Department of Mathematics, School of Applied Mathematics and Physical Sciences, National Technical University of Athens, Zografou Campus, Athens 15780, Greece (chrysafinos@math.ntua.gr).

The aim of the above optimal control problem is to match the velocity vector field to a given target field by influencing the behavior of the system through a control function. The control function is of distributed type and satisfies certain constraints. This is achieved by minimizing the standard tracking-type functional, while the parameter $\lambda > 0$ denotes a penalty parameter, which is typically small compared to the actual size of the data. The terminal term has been included in order to obtain more effective approximations near the end point of the time interval. For related discussions and references regarding the computational significance of the above optimal control problem, we refer the reader to [16].

The analysis of such optimal control problems is well understood. However, when it comes to the approximation and numerical analysis of such problems the existing literature is quite limited. This is due to the fact that the regularity of solutions of Navier–Stokes equations, within the optimal control setting, is very limited, which creates additional difficulties in analyzing suitable schemes for optimal control problems. Standard techniques developed for the numerical analysis of the uncontrolled Navier–Stokes equations cannot be directly applied in the optimal control setting. In addition, optimal control problems constrained to nonlinear evolutionary PDEs with control constraints typically exhibit fine properties and hence require special techniques involving both first and second order necessary and sufficient conditions.

Our work analyzes a numerical scheme based on the discontinuous time-stepping Galerkin scheme for the piecewise constant time combined with standard conforming finite element subspaces for the discretization in space. The main result of our work is to derive space-time error estimates, under suitable regularity assumptions on the data by utilizing ideas from [6] developed for the stationary Navier–Stokes, together with a detailed error analysis of the uncontrolled state and adjoint equations of the underlying scheme. To the best of our knowledge our estimates are new. Two parameters are associated to the numerical discretization, τ and h , indicating the size of the grids in time and space, respectively. The usual assumption $\tau \leq Ch^2$ is needed to prove that the discrete equation has a unique solution. The reader should observe that if we discretize the state equation only in time, not in space, then we cannot prove uniqueness of a solution for the resulting elliptic system. Indeed, this discrete elliptic system is very close to the stationary Navier–Stokes system, for which there is no uniqueness result. Therefore, it is not surprising that the discretization parameter τ needs to be small compared with h if we want to prove the uniqueness of a solution for the fully discrete system. We also make use of this condition to prove error estimates of order $O(h)$. For some related earlier work, we refer the reader to [1], [17], [18], [19], [20], [29], [30], [34] and the reference cited therein. A very close paper is [10]. The main differences with [10] are the nonexistence of control constraints and the fact that the state equation is not discretized in time. The absence of control constraints allows a direct analysis of the system of state and adjoint state equations, which is not possible under control constraints. To overcome this difficulty we need to use the second order conditions for optimality. By using a variational discretization, in [10] the authors can prove error estimates of order $O(h^2)$. The same estimates can be proved when the state equations are fully discretized. The proof of error estimates of order $O(h^2)$ for the variational discretization of the control problem (P) will be the goal of a forthcoming paper currently in preparation.

An interesting reference for the approximation of control problems associated to parabolic semilinear equations is [27]. Neitzel and Vexler discretize the state equation in two steps, first in time and then in space. They take advantage of the boundedness in time-space of the states to get error estimates for the control of order $O(\tau + h)$

without the assumption $\tau \leq Ch^2$. However, they make a strong second order condition that we do not need. Their approach is not easily adapted for the control of Navier–Stokes systems because the nonlinearity involves the gradient of the state, and the boundedness of the states fails. Moreover, the discretization in time of the state equation leads to a stationary Navier–Stokes system, for which we cannot guarantee the uniqueness of a solution.

The discontinuous Galerkin time-stepping schemes are known to perform well in a variety of problems whose solutions satisfy low regularity properties. The discontinuous (in time) Galerkin framework also accommodates many different time-stepping schemes. For example, the lowest order scheme (in time) considered here can be viewed as the implicit Euler scheme, while there is a close similarity between higher order (in time) discontinuous Galerkin schemes and other time-stepping approaches such as Runge–Kutta time-stepping techniques, provided that suitable integration techniques are being used to discretize related integrals (see, e.g., [32]). The key difference between the analysis of the classical implicit Euler scheme and its discontinuous (in time) stepping approach is the use of local (in time) approximation theory tools instead of global (in time) approximation and interpolation tools. In addition, the discontinuous (in time) formulation inherits stability/regularity properties of the underlying PDE, due to its heavily implicit nature. As a result, it leads to an efficient analysis of approximation of problems whose solution satisfies low regularity properties, and in particular to problems where the time derivative is discontinuous, and hence discretization in a completely discontinuous fashion is preferable. On the other hand, continuous (in time) Galerkin schemes typically require much more regularity than the one anticipated from our optimal control problem. For example, the lowest order (in time) continuous (in time) Galerkin scheme corresponds to a Petrov–Galerkin Crank–Nicolson scheme, which requires additional regularity properties even in the case of uncontrolled linear parabolic PDEs (see, e.g., [32]). For earlier work on these schemes within the context of optimal control problems, we refer the reader to [24], [25] for error estimates for an optimal control problem for the heat equation with and without control constraints, respectively, and to [8] for a convergence result for a semilinear parabolic optimal control problem. An analysis of the second order Petrov–Galerkin Crank–Nicolson scheme for an optimal control problem for the heat equation is analyzed in [26] where estimates of second order (in time) are derived. However, the regularity assumptions on the control, state, and adjoint variables are not present in the nonlinear setting of the Navier–Stokes equations. For general results related to discontinuous time-stepping schemes for linear parabolic uncontrolled PDEs, we refer the reader to [11, 12, 13, 14, 32] (see also the references within). Finally, in the recent work of [9], discontinuous time-stepping schemes of arbitrary order for the Navier–Stokes equations in two and three dimensions were examined. Further results concerning the analysis and numerical analysis of the uncontrolled Navier–Stokes equations can be found in the classical works of [15], [21], [22], [31]. For several issues related to the analysis and numerics of optimal control problems, we refer the reader to [33] (see also the references within).

2. Assumptions and preliminary results. Ω is a bounded open and convex subset in \mathbb{R}^2 , Γ being its boundary. The outward unit normal vector to Γ at a point $x \in \Gamma$ is denoted by $\mathbf{n}(x)$. Given $0 < T < +\infty$, we denote $\Omega_T = (0, T) \times \Omega$ and $\Sigma_T = (0, T) \times \Gamma$. We fix the notation for Sobolev spaces: $\mathbf{H}^1(\Omega) = H^1(\Omega; \mathbb{R}^2)$, $\mathbf{H}_0^1(\Omega) = H_0^1(\Omega; \mathbb{R}^2)$, $\mathbf{H}^{-1}(\Omega) = (\mathbf{H}_0^1(\Omega))'$, and $\mathbf{W}^{s,p}(\Omega) = W^{s,p}(\Omega; \mathbb{R}^2)$ for $1 \leq p \leq \infty$.

and $s > 0$. We also consider the spaces of integrable functions,

$$L_0^2(\Omega) = \left\{ w \in L^2(\Omega) : \int_{\Omega} w(x) \, dx = 0 \right\};$$

$\mathbf{L}^p(\Omega) = L^p(\Omega; \mathbb{R}^2)$ and, for a given Banach space X , $L^p(0, T; X)$ will denote the integrable functions defined in $(0, T)$ and taking values in X endowed with the usual norm. Following Lions and Magenes [23, Vol. 1] we put

$$H^{2,1}(\Omega_T) = \left\{ y \in L^2(\Omega_T) : \frac{\partial y}{\partial x_i}, \frac{\partial^2 y}{\partial x_i \partial x_j}, \frac{\partial y}{\partial t} \in L^2(\Omega_T), \, 1 \leq i, j \leq 2 \right\}$$

equipped with the standard norm. In [23, Vol. 1] it is proved that every element of $H^{2,1}(\Omega_T)$, after a modification over a zero measure set, is a continuous function from $[0, T] \rightarrow H^1(\Omega)$. We also set $\mathbf{H}^{2,1}(\Omega_T) = H^{2,1}(\Omega_T) \times H^{2,1}(\Omega_T)$.

We introduce the usual spaces of divergence-free vector fields:

$$\begin{aligned} \mathbf{Y} &= \{ \mathbf{y} \in \mathbf{H}_0^1(\Omega) : \operatorname{div} \mathbf{y} = 0 \text{ in } \Omega \}, \\ \mathbf{H} &= \{ \mathbf{y} \in \mathbf{L}^2(\Omega) : \operatorname{div} \mathbf{y} = 0 \text{ in } \Omega \text{ and } \mathbf{y} \cdot \mathbf{n} = 0 \text{ on } \Gamma \}. \end{aligned}$$

Throughout this paper, we will assume that $\mathbf{f}, \mathbf{u} \in L^2(0, T; \mathbf{L}^2(\Omega))$ and $\mathbf{y}_0 \in \mathbf{Y}$. A solution of (1.1) will be sought in the space $\mathbf{W}(0, T) = \{ \mathbf{y} \in L^2(0, T; \mathbf{Y}) : \mathbf{y}_t \in L^2(0, T; \mathbf{Y}^*) \}$. It is well known that $\mathbf{W}(0, T) \subset C_w([0, T], \mathbf{H})$, where $C_w([0, T], \mathbf{H})$ is the space of weakly continuous functions $\mathbf{y} : [0, T] \rightarrow \mathbf{H}$.

Let us introduce the weak formulation of (1.1). To this end we define the bilinear and trilinear forms $a : \mathbf{H}^1(\Omega) \times \mathbf{H}^1(\Omega) \rightarrow \mathbb{R}$ and $c : \mathbf{L}^4(\Omega) \times \mathbf{H}^1(\Omega) \times \mathbf{H}^1(\Omega) \rightarrow \mathbb{R}$ by

$$\begin{aligned} a(\mathbf{y}, \mathbf{z}) &= \nu \int_{\Omega} (\nabla \mathbf{y} : \nabla \mathbf{z}) \, dx = \nu \sum_{i,j=1}^2 \int_{\Omega} \partial_{x_i} y_j \partial_{x_i} z_j \, dx, \\ c(\mathbf{y}, \mathbf{z}, \mathbf{w}) &= \frac{1}{2} [\hat{c}(\mathbf{y}, \mathbf{z}, \mathbf{w}) - \hat{c}(\mathbf{y}, \mathbf{w}, \mathbf{z})] \text{ with } \hat{c}(\mathbf{y}, \mathbf{z}, \mathbf{w}) = \sum_{i,j=1}^2 \int_{\Omega} y_j \left(\frac{\partial z_i}{\partial x_j} \right) w_i \, dx. \end{aligned}$$

Now, we seek $\mathbf{y} \in \mathbf{W}(0, T)$ such that for a.e. $t \in (0, T)$,

$$(2.1) \quad \begin{cases} (\mathbf{y}_t, \mathbf{w}) + a(\mathbf{y}, \mathbf{w}) + c(\mathbf{y}, \mathbf{y}, \mathbf{w}) = (\mathbf{f} + \mathbf{u}, \mathbf{w}) \quad \forall \mathbf{w} \in \mathbf{Y}, \\ \mathbf{y}(0) = \mathbf{y}_0. \end{cases}$$

Above (\cdot, \cdot) denotes the scalar product in $\mathbf{L}^2(\Omega)$. This notation will be frequently used throughout the paper and $\|\cdot\|$ will denote the associated norm. Any other norm will be indicated by a subscript.

Equation (2.1) has a unique solution in $\mathbf{W}(0, T)$. Once the velocity \mathbf{y} is obtained, then the existence of a pressure $p \in \mathcal{D}(\Omega_T)$ is proved in such a way that the first equation of (1.1) holds in a distribution sense. Thanks to the regularity assumed on \mathbf{f} , \mathbf{y}_0 , and Ω , then some extra regularity is proved for (\mathbf{y}, p) . Indeed, we have that $\mathbf{y} \in \mathbf{H}^{2,1}(\Omega_T) \cap C([0, T], \mathbf{Y})$ and $p \in L^2(0, T; H^1(\Omega))$, the pressure being unique up to an additive constant; see, for instance, Ladyzhenskaya [21], Lions [22], Temam [31].

The next properties of the trilinear form c will be used later. The proof can be found in many books; see [21], [22], or [31].

LEMMA 2.1. *The trilinear form c satisfies*

$$\begin{aligned} c(\mathbf{y}, \mathbf{w}, \mathbf{z}) &= \hat{c}(\mathbf{y}, \mathbf{z}, \mathbf{w}) = -\hat{c}(\mathbf{y}, \mathbf{w}, \mathbf{z}) \quad \forall \mathbf{y} \in \mathbf{Y} \text{ and } \forall \mathbf{z}, \mathbf{w} \in \mathbf{H}_0^1(\Omega), \\ c(\mathbf{y}, \mathbf{z}, \mathbf{w}) &= -c(\mathbf{y}, \mathbf{w}, \mathbf{z}) \quad \forall \mathbf{y} \in \mathbf{L}^4(\Omega) \text{ and } \forall \mathbf{z}, \mathbf{w} \in \mathbf{H}^1(\Omega), \\ c(\mathbf{y}, \mathbf{w}, \mathbf{w}) &= 0 \quad \forall \mathbf{y} \in \mathbf{L}^4(\Omega) \text{ and } \forall \mathbf{w} \in \mathbf{H}^1(\Omega). \end{aligned}$$

Moreover, the following inequalities hold:

$$\begin{aligned} |c(\mathbf{y}, \mathbf{z}, \mathbf{w})| &\leq \|\mathbf{y}\|_{\mathbf{L}^p(\Omega)} \|\nabla \mathbf{z}\|_{\mathbf{L}^2(\Omega)} \|\mathbf{w}\|_{\mathbf{L}^q(\Omega)}, \quad (1/p) + (1/q) = (1/2), \\ |c(\mathbf{y}, \mathbf{z}, \mathbf{w})| &\leq \|\mathbf{y}\|_{\mathbf{L}^4(\Omega)} \|\nabla \mathbf{z}\|_{\mathbf{L}^2(\Omega)} \|\mathbf{w}\|_{\mathbf{L}^4(\Omega)}. \end{aligned}$$

By using the interpolation inequality

$$(2.2) \quad \|z\|_{\mathbf{L}^4(\Omega)} \leq 2^{1/4} \|z\|_{\mathbf{L}^2(\Omega)}^{1/2} \|\nabla z\|_{\mathbf{L}^2(\Omega)}^{1/2} \quad \forall z \in H_0^1(\Omega)$$

(see [31, Lemma 3.3, page 91]), we obtain $\forall \mathbf{y}, \mathbf{w} \in \mathbf{H}_0^1(\Omega)$ and $\forall \mathbf{z} \in \mathbf{H}^1(\Omega)$

$$(2.3) \quad |c(\mathbf{y}, \mathbf{z}, \mathbf{w})| \leq C \|\mathbf{y}\|_{\mathbf{L}^2(\Omega)}^{1/2} \|\nabla \mathbf{y}\|_{\mathbf{L}^2(\Omega)}^{1/2} \|\nabla \mathbf{z}\|_{\mathbf{L}^2(\Omega)} \|\mathbf{w}\|_{\mathbf{L}^2(\Omega)}^{1/2} \|\nabla \mathbf{w}\|_{\mathbf{L}^2(\Omega)}^{1/2}.$$

Returning back to the control problem (P), we will assume

$$(2.4) \quad \lambda > 0, \quad \gamma \geq 0, \quad \mathbf{y}_d \in L^2(0, T; \mathbf{L}^2(\Omega)), \quad \text{and } \mathbf{y}_\Omega \in Y.$$

Since the mapping $G : L^2(0, T; \mathbf{L}^2(\Omega)) \rightarrow \mathbf{H}^{2,1}(\Omega_T) \cap C([0, T]; \mathbf{Y})$, associating to each control \mathbf{u} the corresponding state $G(\mathbf{u}) = \mathbf{y}_\mathbf{u}$ solution of (2.1), is well defined and continuous, then the cost functional $J : L^2(0, T; \mathbf{L}^2(\Omega)) \rightarrow \mathbb{R}$ is also well defined and continuous. The proof of the existence of at least one solution of (P) is standard.

3. Optimality conditions. Since the problem (P) is not convex, we will deal hereafter with global and local solutions. A control $\bar{\mathbf{u}} \in \mathcal{U}_{ad}$ is said to be a local solution of (P) if there exists $\varepsilon > 0$ such that $J(\bar{\mathbf{u}}) \leq J(\mathbf{u})$ for every $\mathbf{u} \in \mathcal{U}_{ad} \cap B_\varepsilon(\bar{\mathbf{u}})$, where $B_\varepsilon(\bar{\mathbf{u}})$ denotes the open ball of $L^2(0, T; \mathbf{L}^2(\Omega))$ centered at $\bar{\mathbf{u}}$ and radius ε . In this section, we establish first and second order optimality conditions for a local solution of problem (P). To this end, we need the differentiability of the mapping G .

THEOREM 3.1 (Casas [4]). *The mapping G is of class C^∞ . Moreover, for any $\mathbf{u}, \mathbf{v} \in L^2(0, T; \mathbf{L}^2(\Omega))$, if we denote $\mathbf{y}_\mathbf{u} = G(\mathbf{u})$, $\mathbf{z}_\mathbf{v} = G'(\mathbf{u})\mathbf{v}$, and $\mathbf{z}_{\mathbf{v}\mathbf{v}} = G''(\mathbf{u})\mathbf{v}^2$, then $\mathbf{z}_\mathbf{v}$ and $\mathbf{z}_{\mathbf{v}\mathbf{v}}$ are the unique solutions of the following equations: $\forall \mathbf{w} \in \mathbf{Y}$*

$$(3.1) \quad \begin{cases} (\mathbf{z}_{\mathbf{v},t}, \mathbf{w}) + a(\mathbf{z}_\mathbf{v}, \mathbf{w}) + c(\mathbf{z}_\mathbf{v}, \mathbf{y}_\mathbf{u}, \mathbf{w}) + c(\mathbf{y}_\mathbf{u}, \mathbf{z}_\mathbf{v}, \mathbf{w}) = (\mathbf{v}, \mathbf{w}), \\ \mathbf{z}_\mathbf{v}(0) = 0, \end{cases}$$

$$(3.2) \quad \begin{cases} (\mathbf{z}_{\mathbf{v}\mathbf{v},t}, \mathbf{w}) + a(\mathbf{z}_{\mathbf{v}\mathbf{v}}, \mathbf{w}) + c(\mathbf{z}_{\mathbf{v}\mathbf{v}}, \mathbf{y}_\mathbf{u}, \mathbf{w}) + c(\mathbf{y}_\mathbf{u}, \mathbf{z}_{\mathbf{v}\mathbf{v}}, \mathbf{w}) + 2c(\mathbf{z}_\mathbf{v}, \mathbf{z}_\mathbf{v}, \mathbf{w}) = 0, \\ \mathbf{z}_{\mathbf{v}\mathbf{v}}(0) = 0. \end{cases}$$

As a consequence of this theorem we get the differentiability of the cost functional.

THEOREM 3.2. *The cost functional $J : L^2(0, T; \mathbf{L}^2(\Omega)) \rightarrow \mathbb{R}$ is of class C^∞ , and for every $\mathbf{u}, \mathbf{v} \in L^2(0, T; \mathbf{L}^2(\Omega))$ we have*

$$\begin{aligned} (3.3) \quad J'(\mathbf{u})\mathbf{v} &= \int_0^T \int_\Omega (\varphi_\mathbf{u} + \lambda \mathbf{u}) \mathbf{v} \, dx dt, \\ J''(\mathbf{u})\mathbf{v}^2 &= \int_0^T \int_\Omega (|\mathbf{z}_\mathbf{v}|^2 - 2(\mathbf{z}_\mathbf{v} \cdot \nabla) \mathbf{z}_\mathbf{v} \varphi_\mathbf{u}) \, dx dt \\ (3.4) \quad &+ \gamma \int_\Omega |\mathbf{z}_\mathbf{v}(T)|^2 \, dx + \lambda \int_0^T \int_\Omega |\mathbf{v}|^2 \, dx dt, \end{aligned}$$

where $\mathbf{z}_v = G'(\mathbf{u})\mathbf{v}$ is the solution of (3.1) and $\varphi_u \in \mathbf{H}^{2,1}(\Omega_T) \cap C([0, T], \mathbf{Y})$ is the unique element satisfying, for every $\mathbf{w} \in \mathbf{Y}$,

$$(3.5) \quad \begin{cases} -(\varphi_{u,t}, \mathbf{w}) + a(\varphi_u, \mathbf{w}) + c(\mathbf{w}, \mathbf{y}_u, \varphi_u) + c(\mathbf{y}_u, \mathbf{w}, \varphi_u) = (\mathbf{y}_u - \mathbf{y}_d, \mathbf{w}), \\ \varphi_u(T) = \gamma(\mathbf{y}(T) - \mathbf{y}_\Omega). \end{cases}$$

Proof. First, let us observe that (3.5) is the adjoint of (3.1). Since (3.1) has a unique solution in $\mathbf{H}^{2,1}(\Omega_T) \cap C([0, T], \mathbf{Y})$ for any $\mathbf{v} \in L^2(0, T; \mathbf{L}^2(\Omega))$, then arguing by transposition we can prove the existence and uniqueness of the solution φ_u of (3.5), as well as the regularity $\varphi_u \in \mathbf{H}^{2,1}(\Omega_T) \cap C([0, T], \mathbf{Y})$. Now, the differentiability property of J and relations (3.3) and (3.4) are a consequence of Theorem 3.1 and the chain rule. \square

Now, we get the optimality conditions. We start with the first order conditions.

THEOREM 3.3. *Let us assume that $\bar{\mathbf{u}}$ is a local solution of problem (P); then there exist $\bar{\mathbf{y}}$ and $\bar{\varphi}$ belonging to $\mathbf{H}^{2,1}(\Omega_T) \cap C([0, T], \mathbf{Y})$ such that*

$$(3.6) \quad \begin{cases} (\bar{\mathbf{y}}_t, \mathbf{w}) + a(\bar{\mathbf{y}}, \mathbf{w}) + c(\bar{\mathbf{y}}, \bar{\mathbf{y}}, \mathbf{w}) = (\mathbf{f} + \bar{\mathbf{u}}, \mathbf{w}) \quad \forall \mathbf{w} \in \mathbf{Y}, \\ \bar{\mathbf{y}}(0) = \mathbf{y}_0, \end{cases}$$

$$(3.7) \quad \begin{cases} -(\bar{\varphi}_t, \mathbf{w}) + a(\bar{\varphi}, \mathbf{w}) + c(\mathbf{w}, \bar{\mathbf{y}}, \bar{\varphi}) + c(\bar{\mathbf{y}}, \mathbf{w}, \bar{\varphi}) = (\bar{\mathbf{y}}_u - \mathbf{y}_d, \mathbf{w}) \quad \forall \mathbf{w} \in \mathbf{Y}, \\ \bar{\varphi}(T) = \gamma(\bar{\mathbf{y}}(T) - \mathbf{y}_\Omega), \end{cases}$$

$$(3.8) \quad \int_0^T \int_\Omega (\bar{\varphi} + \lambda \bar{\mathbf{u}})(\mathbf{u} - \bar{\mathbf{u}}) \, dx dt \geq 0 \quad \forall \mathbf{u} \in \mathcal{U}_{ad}.$$

Moreover, the regularity property $\bar{\mathbf{u}} \in \mathbf{H}^1(\Omega_T) \cap C([0, T], \mathbf{H}^1(\Omega)) \cap L^2(0, T; \mathbf{W}^{1,p}(\Omega))$ holds $\forall 1 \leq p < +\infty$.

Proof. Since \mathcal{U}_{ad} is convex, any local solution $\bar{\mathbf{u}}$ satisfies the condition $J'(\bar{\mathbf{u}})(\mathbf{u} - \bar{\mathbf{u}}) \geq 0$ for every $\mathbf{u} \in \mathcal{U}_{ad}$. Then it is enough to use the expression of the derivative given by (3.3) and take $\bar{\mathbf{y}} = \mathbf{y}_{\bar{\mathbf{u}}}$ and $\bar{\varphi} = \varphi_{\bar{\mathbf{u}}}$ to deduce (3.6)–(3.8). The regularity of $\bar{\mathbf{u}}$ follows from (3.8) as usual; we simply observe that (3.8) implies that

$$(3.9) \quad \bar{u}_j(t, x) = \text{Proj}_{[\alpha_j, \beta_j]} \left(-\frac{1}{\lambda} \bar{\varphi}_j(t, x) \right) \quad \text{for a.a. } (t, x) \in \Omega_T, \quad j = 1, 2. \quad \square$$

To write the second order conditions we need to define the cone of critical directions. To this end, let us introduce the function

$$(3.10) \quad \bar{\mathbf{d}} = \bar{\varphi} + \lambda \bar{\mathbf{u}}.$$

Now we set

$$(3.11) \quad \mathcal{C}_{\bar{\mathbf{u}}} = \{\mathbf{v} \in L^2(0, T; \mathbf{L}^2(\Omega)) : \mathbf{v} \text{ satisfies (3.12)–(3.14)}\},$$

$$(3.12) \quad v_j(t, x) \geq 0 \text{ if } -\infty < \alpha_j = \bar{u}_j(t, x),$$

$$(3.13) \quad v_j(t, x) \leq 0 \text{ if } \bar{u}_j(t, x) = \beta_j < +\infty, \quad j = 1, 2,$$

$$(3.14) \quad v_j(t, x) = 0 \text{ if } \bar{d}_j(t, x) \neq 0.$$

Let us notice that

$$(3.15) \quad \begin{aligned} J'(\bar{\mathbf{u}})\mathbf{v} &= \int_0^T \int_\Omega \bar{\mathbf{d}}(t, x) \cdot \mathbf{v}(t, x) \, dx dt, \\ \bar{\mathbf{d}}(t, x) \cdot \mathbf{v}(t, x) &= 0 \text{ for a.a. } (t, x) \in \Omega_T \text{ and } \forall \mathbf{v} \in \mathcal{C}_{\bar{\mathbf{u}}}. \end{aligned}$$

We also deduce as usual from (3.8), for a.a. $(t, x) \in \Omega_T$ and $j = 1, 2$,

$$(3.16) \quad \begin{cases} \bar{u}_j(t, x) = \alpha_j & \Rightarrow \bar{d}_j(t, x) \geq 0, \\ \bar{u}_j(t, x) = \beta_j & \Rightarrow \bar{d}_j(t, x) \leq 0, \\ \alpha_j < \bar{u}_j(t, x) < \beta_j & \Rightarrow \bar{d}_j(t, x) = 0, \end{cases} \text{ and } \begin{cases} \bar{d}_j(t, x) > 0 \Rightarrow \bar{u}_j(t, x) = \alpha_j, \\ \bar{d}_j(t, x) < 0 \Rightarrow \bar{u}_j(t, x) = \beta_j. \end{cases}$$

THEOREM 3.4. *Let $\bar{\mathbf{u}}$ be a local solution of problem (P); then $J''(\bar{\mathbf{u}})\mathbf{v}^2 \geq 0$ $\forall \mathbf{v} \in \mathcal{C}_{\bar{\mathbf{u}}}$. Conversely, let us assume that $\bar{\mathbf{u}} \in \mathcal{U}_{ad}$ satisfies*

$$(3.17) \quad J'(\bar{\mathbf{u}})(\mathbf{u} - \bar{\mathbf{u}}) \geq 0 \quad \forall \mathbf{u} \in \mathcal{U}_{ad},$$

$$(3.18) \quad J''(\bar{\mathbf{u}})\mathbf{v}^2 > 0 \quad \forall \mathbf{v} \in \mathcal{C}_{\bar{\mathbf{u}}} \setminus \{0\}.$$

Then there exist $\varepsilon > 0$ and $\delta > 0$ such that

$$(3.19) \quad J(\bar{\mathbf{u}}) + \frac{\delta}{2} \|\mathbf{u} - \bar{\mathbf{u}}\|_{L^2(0,T;\mathbf{L}^2(\Omega))}^2 \leq J(\mathbf{u}) \quad \forall \mathbf{u} \in \mathcal{U}_{ad} \cap B_\varepsilon(\bar{\mathbf{u}}),$$

where $B_\varepsilon(\bar{\mathbf{u}})$ is the $L^2(0, T; \mathbf{L}^2(\Omega))$ -ball of center $\bar{\mathbf{u}}$ and radius ε .

The proof of the necessary condition is similar to the one made in [6] for the case of steady-state Navier–Stokes equations. The proof of the sufficient conditions can be obtained arguing by contradiction, analogously to the approach of some previous papers; see, for instance, [2], [5], [6].

Remark 3.5. The gap between the necessary and sufficient optimality conditions given in Theorems 3.3 and 3.4 is minimal—the same as in finite dimensional optimization problems. This problem does not suffer from the typical two-norm discrepancy arising usually in infinite dimensional optimization problems. This is due to the C^2 -differentiability of J with respect to the $L^2(0, T; \mathbf{L}^2(\Omega))$ -norm, thanks to a certain compactness with respect to \mathbf{u} in the first two integrals defining J and the fact that the last one is the square of the norm of the control. On the other hand, it is well known that the condition $J''(\bar{\mathbf{u}})\mathbf{v}^2 > 0$ for every nonzero $\mathbf{v} \neq 0$ belonging to the cone of critical directions is not a sufficient optimality condition, in general, in infinite dimensional optimization problems. An inequality of type $J''(\bar{\mathbf{u}})\mathbf{v}^2 \geq \delta \|\mathbf{v}\|_{L^2(0,T;\mathbf{L}^2(\Omega))}^2$ is required in the infinite dimensional case. In finite dimensions, both conditions are equivalent, but this is not the usual case for infinite dimension. However, in our problem we can prove that both conditions are also equivalent. Indeed, let us observe that (3.19) implies that $\bar{\mathbf{u}}$ is a local solution of the problem

$$(P_\varepsilon) \quad \begin{cases} \min J_\delta(\mathbf{u}) = J(\mathbf{u}) - \frac{\delta}{2} \|\mathbf{u} - \bar{\mathbf{u}}\|_{L^2(0,T;\mathbf{L}^2(\Omega))}^2, \\ \mathbf{u} \in \mathcal{U}_{ad} \cap B_\varepsilon(\bar{\mathbf{u}}). \end{cases}$$

Therefore, from the second order necessary conditions we obtain that $J''_\delta(\bar{\mathbf{u}})\mathbf{v}^2 \geq 0$ for every $\mathbf{v} \in \mathcal{C}_{\bar{\mathbf{u}}}$. It is enough to notice that $J''_\delta(\bar{\mathbf{u}})\mathbf{v}^2 = J''(\bar{\mathbf{u}})\mathbf{v}^2 - \delta \|\mathbf{v}\|_{L^2(0,T;\mathbf{L}^2(\Omega))}^2$ to conclude that (3.17)–(3.18) imply

$$(3.20) \quad J''(\bar{\mathbf{u}})\mathbf{v}^2 \geq \delta \|\mathbf{v}\|^2 \quad \forall \mathbf{v} \in \mathcal{C}_{\bar{\mathbf{u}}}.$$

4. Numerical approximation of the control problem. In this section we consider the complete discretization of the control problem (P). To this end, we consider a family of triangulations $\{\mathcal{T}_h\}_{h>0}$ of $\bar{\Omega}$, defined in the standard way, e.g., in [3, Chapter 3.3]. With each element $T \in \mathcal{T}_h$, we associate two parameters h_T and

ϱ_T , where h_T denotes the diameter of the set T and ϱ_T is the diameter of the largest ball contained in T . Define the size of the mesh by $h = \max_{T \in \mathcal{T}_h} h_T$. We also assume that the following regularity assumptions on the triangulation are satisfied:

(i) There exist two positive constants $\varrho_{\mathcal{T}}$ and $\delta_{\mathcal{T}}$ such that $\frac{h_T}{\varrho_T} \leq \varrho_{\mathcal{T}}$ and $\frac{h}{h_T} \leq \delta_{\mathcal{T}}$ $\forall T \in \mathcal{T}_h$ and $\forall h > 0$.

(ii) Define $\overline{\Omega}_h = \cup_{T \in \mathcal{T}_h} T$, and let Ω_h and Γ_h denote its interior and its boundary, respectively. We assume that the vertices of \mathcal{T}_h placed on the boundary Γ_h are points of Γ .

Since Ω is convex, from the last assumption we have that Ω_h is also convex. Moreover, we know that

$$(4.1) \quad |\Omega \setminus \Omega_h| \leq Ch^2;$$

see, for instance, [28, estimate (5.2.19)].

On the mesh \mathcal{T}_h we consider two finite dimensional spaces $\mathbf{Z}_h \subset \mathbf{H}_0^1(\Omega)$ and $Q_h \subset L_0^2(\Omega)$ formed by piecewise polynomials in Ω_h and vanishing in $\Omega \setminus \Omega_h$. We make the following assumptions on these spaces.

(A1) If $\mathbf{z} \in \mathbf{H}^{1+l}(\Omega) \cap \mathbf{H}_0^1(\Omega)$, then

$$(4.2) \quad \inf_{\mathbf{z}_h \in \mathbf{Z}_h} \|\mathbf{z} - \mathbf{z}_h\|_{\mathbf{H}^s(\Omega_h)} \leq Ch^{l+1-s} \|\mathbf{z}\|_{\mathbf{H}^{1+l}(\Omega)} \quad \text{for } 0 \leq l \leq 1 \text{ and } s = 0, 1.$$

(A2) If $q \in H^l(\Omega) \cap L_0^2(\Omega)$, then

$$(4.3) \quad \inf_{q_h \in Q_h} \|q - q_h\|_{L^2(\Omega_h)} \leq Ch \|q\|_{H^1(\Omega)}.$$

(A3) The subspaces \mathbf{Z}_h and Q_h satisfy the inf-sup condition: $\exists c > 0$ such that

$$(4.4) \quad \inf_{q_h \in Q_h} \sup_{\mathbf{z}_h \in \mathbf{Z}_h} \frac{b(\mathbf{z}_h, q_h)}{\|\mathbf{z}_h\|_{\mathbf{H}^1(\Omega_h)} \|q_h\|_{L^2(\Omega_h)}} \geq c,$$

where $b : \mathbf{H}^1(\Omega) \times L^2(\Omega) \rightarrow \mathbb{R}$ is defined by

$$b(\mathbf{z}, q) = \int_{\Omega} q(x) \operatorname{div} \mathbf{z}(x) \, dx.$$

These assumptions are satisfied by the usual finite elements considered in the discretization of Navier–Stokes equations: “Taylor–Hood,” P1-bubble finite element, and some others; see [15, Chapter 2].

We also consider a subspace \mathbf{Y}_h of \mathbf{Z}_h defined by

$$\mathbf{Y}_h = \{\mathbf{y}_h \in \mathbf{Z}_h : b(\mathbf{y}_h, q_h) = 0 \, \forall q_h \in Q_h\}$$

and set

$$\mathbf{U}_h = \{\mathbf{u}_h \in \mathbf{L}^2(\Omega_h) : \mathbf{u}_h|_{\Gamma} \equiv \mathbf{u}_T \in \mathbb{R}^2\}.$$

We proceed now with the discretization in time. Let us consider a grid of points $0 = t_0 < t_1 < \dots < t_{N_\tau} = T$. We denote $\tau_n = t_n - t_{n-1}$. We make the following assumption:

$$(4.5) \quad \exists \varrho_0 > 0 \text{ such that } \tau = \max_{1 \leq n \leq N_\tau} \tau_n < \varrho_0 \tau_n \, \forall 1 \leq n \leq N_\tau \text{ and } \forall \tau > 0.$$

Given a triangulation \mathcal{T}_h of Ω and a grid of points $\{t_n\}_{n=0}^{N_\tau}$ of $[0, T]$, we set $\sigma = (\tau, h)$. Finally, we consider the spaces

$$\begin{aligned}\mathcal{Y}_\sigma &= \{\mathbf{y}_\sigma \in L^2(0, T; \mathbf{Y}_h) : \mathbf{y}_\sigma|_{(t_{n-1}, t_n)} \in \mathbf{Y}_h \text{ for } 1 \leq n \leq N_\tau\}, \\ \mathcal{Q}_\sigma &= \{q_\sigma \in L^2(0, T; Q_h) : q_\sigma|_{(t_{n-1}, t_n)} \in Q_h \text{ for } 1 \leq n \leq N_\tau\}, \\ \mathcal{U}_\sigma &= \{\mathbf{u}_\sigma \in L^2(0, T; \mathbf{U}_h) : \mathbf{u}_\sigma|_{(t_{n-1}, t_n)} \in \mathbf{U}_h \text{ for } 1 \leq n \leq N_\tau\}.\end{aligned}$$

We have that the functions of \mathcal{Y}_σ , \mathcal{Q}_σ , and \mathcal{U}_σ are piecewise constant in time. We will look for the discrete controls in the space \mathcal{U}_σ . An element of this space can be written in the form

$$(4.6) \quad \mathbf{u}_\sigma = \sum_{n=1}^{N_\tau} \sum_{T \in \mathcal{T}_h} \mathbf{u}_{n,T} \chi_n \chi_T, \quad \text{with } \mathbf{u}_{n,T} \in \mathbb{R}^2,$$

where χ_n and χ_T are the characteristic functions of (t_{n-1}, t_n) and T , respectively. Therefore, the dimension of \mathcal{U}_σ is $2N_\tau N_h$, where N_h is the number of triangles in \mathcal{T}_h . In \mathcal{U}_σ we consider the convex subset

$$\mathcal{U}_{\sigma,ad} = \mathcal{U}_\sigma \cap \mathcal{U}_{ad} = \{\mathbf{u}_\sigma \in \mathcal{U}_\sigma : \mathbf{u}_{n,T} \in [\alpha_1, \beta_1] \times [\alpha_2, \beta_2]\}.$$

On the other hand, the elements of \mathcal{Y}_σ can be written in the form

$$(4.7) \quad \mathbf{y}_\sigma = \sum_{n=1}^{N_\tau} \mathbf{y}_{n,h} \chi_n, \quad \text{with } \mathbf{y}_{n,h} \in \mathbf{Y}_h,$$

where χ_n is as above. For every discrete state \mathbf{y}_σ we will fix $\mathbf{y}_\sigma(t_n) = \mathbf{y}_{n,h}$, so that \mathbf{y}_σ is continuous on the left. In particular, we have $\mathbf{y}_\sigma(T) = \mathbf{y}_\sigma(t_{N_\tau}) = \mathbf{y}_{N_\tau,h}$.

To define the discrete control problem we have to consider the numerical discretization of the state equation (1.1), or equivalently (2.1). We achieve this goal by using a discontinuous time-stepping Galerkin method, with piecewise constants in time and conforming finite element spaces in space. For any $\mathbf{u} \in L^2(0, T; \mathbf{L}^2(\Omega))$ the discrete state equation is given by the following:

$$(4.8) \quad \begin{cases} \text{For } n = 1, \dots, N_\tau \text{ and } \forall \mathbf{w}_h \in \mathbf{Y}_h, \\ \left(\frac{\mathbf{y}_{n,h} - \mathbf{y}_{n-1,h}}{\tau_n}, \mathbf{w}_h \right) + a(\mathbf{y}_{n,h}, \mathbf{w}_h) + c(\mathbf{y}_{n,h}, \mathbf{y}_{n,h}, \mathbf{w}_h) = (\mathbf{f}_n + \mathbf{u}_n, \mathbf{w}_h), \\ \mathbf{y}_{0,h} = \mathbf{y}_{0h}, \end{cases}$$

where

$$(4.9) \quad (\mathbf{f}_n, \mathbf{w}_h) = \frac{1}{\tau_n} \int_{t_{n-1}}^{t_n} (\mathbf{f}(t), \mathbf{w}_h) dt, \quad (\mathbf{u}_n, \mathbf{w}_h) = \frac{1}{\tau_n} \int_{t_{n-1}}^{t_n} (\mathbf{u}(t), \mathbf{w}_h) dt,$$

$$(4.10) \quad \mathbf{y}_{0h} \in \mathbf{Y}_h \text{ with } \|\mathbf{y}_0 - \mathbf{y}_{0h}\|_{\mathbf{L}^2(\Omega_h)} \leq Ch \text{ and } \|\mathbf{y}_{0h}\|_{\mathbf{H}^1(\Omega_h)} \leq C \quad \forall h > 0.$$

The above scheme is essentially an implicit Euler in time/conforming in space scheme, and can be easily extended to higher order polynomial in time discretizations; see, e.g., [32] and the references within. For stability and error estimates under suitable regularity assumptions for high order discontinuous time-stepping schemes, we refer the reader to [9]. Here, we focus on the lowest case of polynomial approximation

in time, due to the low regularity imposed by the nature of our optimal control problem. A key feature of the proposed scheme is that the regularity properties of the discrete solution mimics the continuous problem. We will prove later that for any $\mathbf{u} \in L^2(0, T; \mathbf{L}^2(\Omega))$, (4.8) has a unique solution $\mathbf{y}_\sigma(\mathbf{u}) \in \mathcal{Y}_\sigma$. Then we can define the discrete control problem as follows:

$$(P_\sigma) \quad \begin{cases} \min J_\sigma(\mathbf{u}_\sigma) \\ \mathbf{u}_\sigma \in \mathcal{U}_{\sigma, ad}, \end{cases}$$

$$J_\sigma(\mathbf{u}_\sigma) = \frac{1}{2} \int_0^T \int_{\Omega_h} |\mathbf{y}_\sigma(\mathbf{u}_\sigma) - \mathbf{y}_d|^2 dx dt$$

$$+ \frac{\gamma}{2} \int_{\Omega_h} |\mathbf{y}_\sigma(T) - \mathbf{y}_{\Omega_h}|^2 dx + \frac{\lambda}{2} \int_0^T \int_{\Omega_h} |\mathbf{u}_\sigma|^2 dx dt,$$

$$(4.11) \quad \mathbf{y}_{\Omega_h} \in \mathbf{Y}_h \quad \text{with} \quad \|\mathbf{y}_\Omega - \mathbf{y}_{\Omega_h}\|_{\mathbf{L}^2(\Omega_h)} \leq Ch \quad \text{and} \quad \|\mathbf{y}_{\Omega_h}\|_{\mathbf{H}^1(\Omega_h)} \leq C \quad \forall h > 0.$$

In the study of the control problem, first we analyze the discrete state equation (4.8); then we study the discrete adjoint state equation; next we prove the convergence of (P_σ) ; and finally we prove the error estimates for the discretization.

4.1. Analysis of the discrete state equation. By a standard argument, using the identity $c(\mathbf{z}, \mathbf{w}, \mathbf{w}) = 0 \quad \forall \mathbf{z} \in \mathbf{L}^4(\Omega)$ and $\forall \mathbf{w} \in \mathbf{H}^1(\Omega)$ (Lemma 2.1) and Brower's fixed-point theorem, we can easily prove that (4.8) has at least one solution. In this section, we will prove that the solution is unique under some restrictions on $\sigma = (\tau, h)$. For the moment, let us denote $\mathbf{y} = \mathbf{y}_\mathbf{u} = G(\mathbf{u})$ and $\mathbf{y}_\sigma \in \mathcal{Y}_\sigma$ a solution of (4.8). We are going to prove some error estimates for $\mathbf{y} - \mathbf{y}_\sigma$. To this end, we need to introduce some projection operators.

DEFINITION 4.1. We define the projection operator $P_h : \mathbf{L}^2(\Omega) \longrightarrow \mathbf{Y}_h$ by

$$(P_h \mathbf{y}, \mathbf{w}_h) = (\mathbf{y}, \mathbf{w}_h) \quad \forall \mathbf{w}_h \in \mathbf{Y}_h.$$

We also define $P_\sigma : C([0, T], \mathbf{L}^2(\Omega)) \longrightarrow \mathcal{Y}_\sigma$ by $(P_\sigma \mathbf{y})_{n,h} = P_h \mathbf{y}(t_n)$ for $1 \leq n \leq N_\tau$.

LEMMA 4.2. There exists a constant $C > 0$ independent of σ such that for every $\mathbf{y} \in \mathbf{H}^{2,1}(\Omega_T) \cap C([0, T]; \mathbf{Y})$ the following estimate holds:

$$(4.12) \quad \|\mathbf{y} - P_\sigma \mathbf{y}\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))} \leq C \left\{ \tau \|\mathbf{y}'\|_{L^2(0,T;\mathbf{L}^2(\Omega))} + h^2 \|\mathbf{y}\|_{L^2(0,T;\mathbf{H}^2(\Omega))} \right\}.$$

Proof. From assumptions (A1)–(A3) and using (4.2) with $s = 0$ and $l = 1$ (see also [15, Chapter II]), the definition of P_σ , and the stability of P_h we get

$$\begin{aligned} \|\mathbf{y} - P_\sigma \mathbf{y}\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))} &= \left\{ \sum_{n=1}^{N_\tau} \int_{t_{n-1}}^{t_n} \|\mathbf{y}(t) - P_h \mathbf{y}(t_n)\|^2 dt \right\}^{1/2} \\ &\leq \left\{ \sum_{n=1}^{N_\tau} \int_{t_{n-1}}^{t_n} \|\mathbf{y}(t) - P_h \mathbf{y}(t)\|^2 dt \right\}^{1/2} + \left\{ \sum_{n=1}^{N_\tau} \int_{t_{n-1}}^{t_n} \|P_h \mathbf{y}(t) - P_h \mathbf{y}(t_n)\|^2 dt \right\}^{1/2} \\ &\leq Ch^2 \left\{ \sum_{n=1}^{N_\tau} \int_{t_{n-1}}^{t_n} \|\mathbf{y}(t)\|_{\mathbf{H}^2(\Omega)}^2 dt \right\}^{1/2} + \left\{ \sum_{n=1}^{N_\tau} \int_{t_{n-1}}^{t_n} \|\mathbf{y}(t) - \mathbf{y}(t_n)\|^2 dt \right\}^{1/2} \\ &\leq Ch^2 \|\mathbf{y}\|_{L^2(0,T;\mathbf{H}^2(\Omega))} + \left\{ \sum_{n=1}^{N_\tau} \int_{t_{n-1}}^{t_n} (t_n - t) \int_{t_{n-1}}^{t_n} \|\mathbf{y}'(s)\|^2 ds dt \right\}^{1/2} \\ &\leq C \left\{ h^2 \|\mathbf{y}\|_{L^2(0,T;\mathbf{H}^2(\Omega))} + \tau \|\mathbf{y}'\|_{L^2(0,T;\mathbf{L}^2(\Omega))} \right\}. \quad \square \end{aligned}$$

DEFINITION 4.3. The operator $\Pi_h : \mathbf{Y} \rightarrow \mathbf{Y}_h$ is defined by

$$a(\Pi_h \mathbf{y}, \mathbf{w}_h) = a(\mathbf{y}, \mathbf{w}_h) \quad \forall \mathbf{w}_h \in \mathbf{Y}_h.$$

To any element $\mathbf{y} \in C([0, T]; \mathbf{Y})$, we associate $\mathbf{y}_h \in C([0, T], \mathbf{Y}_h)$ by $\mathbf{y}_h(t) = \Pi_h(\mathbf{y}(t))$.

The next lemma is an immediate consequence of assumptions (A1)–(A3); see again [15, Chapter II].

LEMMA 4.4. There exists a constant $C > 0$ independent of h such that

$$(4.13) \quad \|\mathbf{y} - \Pi_h \mathbf{y}\|_{\mathbf{H}^s(\Omega_h)} \leq Ch^{2-s} \|\mathbf{y}\|_{\mathbf{H}^2(\Omega)} \quad \forall \mathbf{y} \in \mathbf{H}^2(\Omega) \cap \mathbf{Y} \quad \text{and} \quad s = 0, 1.$$

As a consequence of the previous two lemmas we have the following result.

LEMMA 4.5. There exists a constant $C > 0$ independent of σ such that for every $\mathbf{y} \in \mathbf{H}^{2,1}(\Omega_T) \cap C([0, T]; \mathbf{Y})$

$$(4.14) \quad \|\mathbf{y} - P_\sigma \mathbf{y}\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))} \leq C \left\{ \frac{\tau}{h} \|\mathbf{y}'\|_{L^2(0,T;\mathbf{L}^2(\Omega))} + h \|\mathbf{y}\|_{L^2(0,T;\mathbf{H}^2(\Omega))} \right\}.$$

Proof. From Lemma 4.4 we get

$$\begin{aligned} \|\mathbf{y} - P_\sigma \mathbf{y}\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))} &\leq \|\mathbf{y} - \mathbf{y}_h\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))} + \|\mathbf{y}_h - P_\sigma \mathbf{y}\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))} \\ &\leq Ch \|\mathbf{y}\|_{L^2(0,T;\mathbf{H}^2(\Omega))} + \|\mathbf{y}_h - P_\sigma \mathbf{y}\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))}. \end{aligned}$$

Now, using the definition of P_σ , an inverse inequality, (4.12), and (4.13) we obtain

$$\begin{aligned} \|\mathbf{y}_h - P_\sigma \mathbf{y}\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))} &= \left\{ \sum_{n=1}^{N_\tau} \int_{t_{n-1}}^{t_n} \|\mathbf{y}_h(t) - P_h \mathbf{y}(t_n)\|_{\mathbf{H}^1(\Omega_h)}^2 dt \right\}^{1/2} \\ &\leq \frac{C}{h} \left\{ \sum_{n=1}^{N_\tau} \int_{t_{n-1}}^{t_n} \|\mathbf{y}_h(t) - P_h \mathbf{y}(t_n)\|_{\mathbf{L}^2(\Omega_h)}^2 dt \right\}^{1/2} \\ &\leq \frac{C}{h} \left\{ \|\mathbf{y} - \mathbf{y}_h\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))} + \|\mathbf{y} - P_\sigma \mathbf{y}\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))} \right\} \\ &\leq C \left\{ \frac{\tau}{h} \|\mathbf{y}'\|_{L^2(0,T;\mathbf{L}^2(\Omega))} + h \|\mathbf{y}\|_{L^2(0,T;\mathbf{H}^2(\Omega))} \right\}. \quad \square \end{aligned}$$

Before proving the error estimates for $\mathbf{y} - \mathbf{y}_\sigma$, we need to establish the corresponding estimates for the Stokes problem. Let us formulate this result as follows.

LEMMA 4.6. Let $\mathbf{y} \in \mathbf{H}^{2,1}(\Omega_T) \cap C([0, T], \mathbf{Y})$ be the solution of (2.1) and let $\hat{\mathbf{y}}_\sigma \in \mathcal{Y}_\sigma$ satisfy the following:

$$(4.15) \quad \begin{cases} \text{For } n = 1, \dots, N_\tau \quad \text{and} \quad \forall \mathbf{w}_h \in \mathbf{Y}_h, \\ \left(\frac{\hat{\mathbf{y}}_{n,h} - \hat{\mathbf{y}}_{n-1,h}}{\tau_n}, \mathbf{w}_h \right) + a(\hat{\mathbf{y}}_{n,h}, \mathbf{w}_h) = (\hat{\mathbf{f}}_n, \mathbf{w}_h), \\ \hat{\mathbf{y}}_{0,h} = \mathbf{y}_{0h}, \end{cases}$$

where $(\hat{\mathbf{f}}_n, \mathbf{w}_h) = \frac{1}{\tau_n} \int_{t_{n-1}}^{t_n} \{a(\mathbf{y}(t), \mathbf{w}_h) + (\mathbf{y}'(t), \mathbf{w}_h)\} dt$. Then (4.15) has a unique solution $\hat{\mathbf{y}}_\sigma \in \mathcal{Y}_\sigma$. Moreover, the following properties hold:

- (1) $\{\hat{\mathbf{y}}_\sigma\}_\sigma$ is bounded in $L^\infty(0, T; \mathbf{H}^1(\Omega_h))$.

(2) *There exists a constant $C > 0$ independent of σ such that*

$$(4.16) \quad \max_{1 \leq n \leq N_\tau} \|\mathbf{y}(t_n) - \hat{\mathbf{y}}_\sigma(t_n)\| + \|\mathbf{y} - \hat{\mathbf{y}}_\sigma\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))} \\ \leq C \left\{ \frac{\tau}{h} \|\mathbf{y}'\|_{L^2(0,T;\mathbf{L}^2(\Omega))} + h \|\mathbf{y}\|_{L^2(0,T;\mathbf{H}^2(\Omega))} + h \|\mathbf{y}_0\|_{\mathbf{H}^1(\Omega)} \right\},$$

$$(4.17) \quad \|\mathbf{y} - \hat{\mathbf{y}}_\sigma\|_{L^\infty(0,T;\mathbf{L}^2(\Omega_h))} \\ \leq C \left\{ \left(\frac{\tau}{h} + \sqrt{\tau} \right) \|\mathbf{y}'\|_{L^2(0,T;\mathbf{L}^2(\Omega))} + h \|\mathbf{y}\|_{L^2(0,T;\mathbf{H}^2(\Omega))} + h \|\mathbf{y}_0\|_{\mathbf{H}^1(\Omega)} \right\}.$$

Proof. The existence and uniqueness of the solution $\hat{\mathbf{y}}_\sigma$ is easy and well known. The boundedness of $\{\hat{\mathbf{y}}_\sigma\}_\sigma$ in $L^\infty(0,T;\mathbf{H}^1(\Omega_h))$ was proved in [9, Theorem 4.10]. The estimate (4.16) follows from (4.14) and [9, Theorem 4.6]. Finally, we prove (4.17). Let us assume that $t_{n-1} < t < t_n$ for some $1 \leq N_\tau$. Then

$$\|\mathbf{y}(t) - \hat{\mathbf{y}}_\sigma(t)\| \leq \|\mathbf{y}(t) - \mathbf{y}(t_n)\| + \|\mathbf{y}(t_n) - \hat{\mathbf{y}}_\sigma(t_n)\|.$$

The second term on the right-hand side of the inequality has been estimated in (4.16). Let us study the first term. For any $\mathbf{w} \in \mathbf{L}^2(\Omega)$

$$|(\mathbf{y}(t) - \mathbf{y}(t_n), \mathbf{w})| = \left| \int_t^{t_n} (\mathbf{y}'(s), \mathbf{w}) \, ds \right| \leq \int_t^{t_n} \|\mathbf{y}'(s)\| \, ds \|\mathbf{w}\| \\ \leq \sqrt{\tau} \|\mathbf{y}'\|_{L^2(0,T;\mathbf{L}^2(\Omega))} \|\mathbf{w}\|;$$

hence $\|\mathbf{y}(t) - \mathbf{y}(t_n)\| \leq \sqrt{\tau} \|\mathbf{y}'\|_{L^2(0,T;\mathbf{L}^2(\Omega))}$. This estimate and (4.16) infer (4.17). \square

The discrete solution of the linear Stokes problem will subsequently play the role of a global-in-time projection, which facilitates the derivation of error estimates under the restricted regularity assumptions of the control problem (see also [9]). Finally, we obtain the result concerning the discrete state equation (4.8).

THEOREM 4.7. *Given $\mathbf{u} \in L^2(0,T;\mathbf{L}^2(\Omega))$, let $\mathbf{y} \in \mathbf{H}^{2,1}(\Omega_T) \cap C([0,T];\mathbf{Y})$ be the solution of (2.1) and let $\mathbf{y}_\sigma \in \mathcal{Y}_\sigma$ be any solution of (4.8). Then there exists a constant $C > 0$ independent of \mathbf{u} , \mathbf{y} , and σ such that*

$$(4.18) \quad \max_{1 \leq n \leq N_\tau} \|\mathbf{y}(t_n) - \mathbf{y}_\sigma(t_n)\| + \|\mathbf{y} - \mathbf{y}_\sigma\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))} \\ \leq C \left\{ \frac{\tau}{h} \|\mathbf{y}'\|_{L^2(0,T;\mathbf{L}^2(\Omega))} + h \|\mathbf{y}\|_{L^2(0,T;\mathbf{H}^2(\Omega))} + h \|\mathbf{y}_0\|_{\mathbf{H}^1(\Omega)} \right\},$$

$$(4.19) \quad \|\mathbf{y} - \mathbf{y}_\sigma\|_{L^\infty(0,T;\mathbf{L}^2(\Omega_h))} \\ \leq C \left\{ \left(\frac{\tau}{h} + \sqrt{\tau} \right) \|\mathbf{y}'\|_{L^2(0,T;\mathbf{L}^2(\Omega))} + h \|\mathbf{y}\|_{L^2(0,T;\mathbf{H}^2(\Omega))} + h \|\mathbf{y}_0\|_{\mathbf{H}^1(\Omega)} \right\}.$$

Moreover, if there exists a constant $C_0 > 0$ such that $\tau \leq C_0 h^2$ for every $\sigma = (\tau, h)$, then $\{\mathbf{y}_\sigma\}_\sigma$ is bounded in $L^\infty(0,T;\mathbf{H}^1(\Omega_h))$ and (4.8) has a unique solution.

Proof. Let us define $\mathbf{e} = \mathbf{y} - \mathbf{y}_\sigma = (\mathbf{y} - \hat{\mathbf{y}}_\sigma) + (\hat{\mathbf{y}}_\sigma - \mathbf{y}_\sigma) = \hat{\mathbf{e}} + \mathbf{e}_\sigma$, where $\hat{\mathbf{y}}_\sigma$ is the solution of (4.15). Then we can proceed as in [9, Theorem 5.2] to get

$$(4.20) \quad \frac{1}{2} \|\mathbf{e}_{n,h}\|^2 - \frac{1}{2} \|\mathbf{e}_{n-1,h}\|^2 + \frac{1}{2} \|\mathbf{e}_{n,h} - \mathbf{e}_{n-1,h}\|^2 + \nu \int_{t_{n-1}}^{t_n} \|\nabla \mathbf{e}_{n,h}\|^2 \, dt \\ \leq \int_{t_{n-1}}^{t_n} \{ |c(\mathbf{e}_{n,h}, \hat{\mathbf{y}}_{n,h}, \mathbf{e}_{n,h})| + |c(\hat{\mathbf{e}}(t), \mathbf{y}(t), \mathbf{e}_{n,h})| + |c(\hat{\mathbf{y}}_{n,h}, \hat{\mathbf{e}}(t), \mathbf{e}_{n,h})| \} \, dt.$$

It remains to estimate the last three terms. For the first we use that $\{\hat{\mathbf{y}}_\sigma\}_\sigma$ is bounded in $L^\infty(0, T; \mathbf{H}^1(\Omega_h))$ (see Lemma 4.6) and (2.2). Then

$$\begin{aligned} \int_{t_{n-1}}^{t_n} |c(\mathbf{e}_{n,h}, \hat{\mathbf{y}}_{n,h}, \mathbf{e}_{n,h})| dt &\leq C \int_{t_{n-1}}^{t_n} \|\mathbf{e}_{n,h}\| \|\nabla \mathbf{e}_{n,h}\| dt \\ &\leq \frac{C\tau_n}{\nu} \|\mathbf{e}_{n,h}\|^2 + \frac{\nu}{4} \int_{t_{n-1}}^{t_n} \|\nabla \mathbf{e}_{n,h}\|^2 dt. \end{aligned}$$

For the second term we use that $\mathbf{y} \in L^\infty(0, T; \mathbf{H}^1(\Omega))$,

$$\begin{aligned} \int_{t_{n-1}}^{t_n} |c(\hat{\mathbf{e}}(t), \mathbf{y}(t), \mathbf{e}_{n,h})| dt &\leq C \int_{t_{n-1}}^{t_n} \|\nabla \hat{\mathbf{e}}(t)\| \|\nabla \mathbf{e}_{n,h}\| dt \\ &\leq \frac{C}{\nu} \int_{t_{n-1}}^{t_n} \|\nabla \hat{\mathbf{e}}(t)\|^2 dt + \frac{\nu}{4} \int_{t_{n-1}}^{t_n} \|\nabla \mathbf{e}_{n,h}\|^2 dt. \end{aligned}$$

Finally, using again the boundedness of $\{\hat{\mathbf{y}}_\sigma\}_\sigma$ in $L^\infty(0, T; \mathbf{H}^1(\Omega_h))$, we get the same estimate as the last one for the third term. Putting all these estimates in (4.20) we obtain

$$\begin{aligned} (1 - C\tau_n) \|\mathbf{e}_{n,h}\|^2 + \|\mathbf{e}_{n,h} - \mathbf{e}_{n-1,h}\|^2 + \frac{\nu}{2} \int_{t_{n-1}}^{t_n} \|\nabla \mathbf{e}_{n,h}\|^2 dt \\ \leq \|\mathbf{e}_{n-1,h}\|^2 + C \|\hat{\mathbf{e}}\|_{L^2(t_{n-1}, t_n; \mathbf{H}^1(\Omega))}^2. \end{aligned}$$

Then, using the discrete Gronwall inequality and the fact that $\mathbf{e}_{0,h} = 0$, we get

$$\|\mathbf{e}_{n,h}\|^2 + \frac{\nu}{2} \int_0^{t_n} \|\nabla \mathbf{e}_{n,h}\|^2 dt \leq C \|\hat{\mathbf{e}}\|_{L^2(0, T; \mathbf{H}^1(\Omega))}^2 \quad \forall 1 \leq n \leq N_\tau.$$

This inequality along with (4.16) and the identity $\mathbf{y} - \mathbf{y}_\sigma = \hat{\mathbf{e}} + \mathbf{e}_\sigma$ proves (4.18). Arguing as in the proof of (4.17), we deduce (4.19) from (4.18). The proof of the boundedness of $\{\mathbf{y}_\sigma\}_\sigma$ in $L^\infty(0, T; \mathbf{H}^1(\Omega_h))$ is an easy consequence of the previous results. Indeed, first we recall that $\{\hat{\mathbf{y}}_\sigma\}_\sigma$ is bounded in $L^\infty(0, T; \mathbf{H}^1(\Omega_h))$ (Lemma 4.6). Now, we write

$$\|\mathbf{y}_\sigma\|_{L^\infty(0, T; \mathbf{H}^1(\Omega_h))} \leq \|\mathbf{y}_\sigma - \hat{\mathbf{y}}_\sigma\|_{L^\infty(0, T; \mathbf{H}^1(\Omega_h))} + \|\hat{\mathbf{y}}_\sigma\|_{L^\infty(0, T; \mathbf{H}^1(\Omega_h))}.$$

It is enough to prove the boundedness of the first term. From an inverse inequality [3, section 4.5], the estimates (4.17) and (4.19), and the inequality $\tau \leq C_0 h^2$ we get

$$\begin{aligned} \|\mathbf{y}_\sigma - \hat{\mathbf{y}}_\sigma\|_{L^\infty(0, T; \mathbf{H}^1(\Omega_h))} &\leq \frac{C}{h} \|\mathbf{y}_\sigma - \hat{\mathbf{y}}_\sigma\|_{L^\infty(0, T; \mathbf{L}^2(\Omega_h))} \\ &\leq \frac{C}{h} \{ \|\mathbf{y}_\sigma - \mathbf{y}\|_{L^\infty(0, T; \mathbf{L}^2(\Omega_h))} + \|\mathbf{y} - \hat{\mathbf{y}}_\sigma\|_{L^\infty(0, T; \mathbf{L}^2(\Omega_h))} \} \leq C \quad \forall \sigma. \end{aligned}$$

To conclude the proof, we have to show the uniqueness of a solution of (4.8). Let us assume that $\mathbf{y}_\sigma^1, \mathbf{y}_\sigma^2 \in \mathcal{Y}_\sigma$ are two solutions of (4.8). Then we set $\mathbf{y}_\sigma = \mathbf{y}_\sigma^2 - \mathbf{y}_\sigma^1$ and we will prove that $\mathbf{y}_\sigma = 0$. Subtracting (4.8) for \mathbf{y}_σ^2 and \mathbf{y}_σ^1 and setting $\mathbf{w}_h = \mathbf{y}_{n,h}$ we get

$$\left(\frac{\mathbf{y}_{n,h} - \mathbf{y}_{n-1,h}}{\tau_n}, \mathbf{y}_{n,h} \right) + a(\mathbf{y}_{n,h}, \mathbf{y}_{n,h}) = c(\mathbf{y}_{n,h}^1, \mathbf{y}_{n,h}^1, \mathbf{y}_{n,h}) - c(\mathbf{y}_{n,h}^2, \mathbf{y}_{n,h}^2, \mathbf{y}_{n,h}).$$

Since $c(\mathbf{y}_{n,h}^2, \mathbf{y}_{n,h}, \mathbf{y}_{n,h}) = 0$, then $c(\mathbf{y}_{n,h}^2, \mathbf{y}_{n,h}^2, \mathbf{y}_{n,h}) = c(\mathbf{y}_{n,h}^2, \mathbf{y}_{n,h}^1, \mathbf{y}_{n,h})$, and therefore

$$c(\mathbf{y}_{n,h}^1, \mathbf{y}_{n,h}^1, \mathbf{y}_{n,h}) - c(\mathbf{y}_{n,h}^2, \mathbf{y}_{n,h}^2, \mathbf{y}_{n,h}) = -c(\mathbf{y}_{n,h}, \mathbf{y}_{n,h}^1, \mathbf{y}_{n,h}).$$

Using this in the above identity and the boundedness of $\{\mathbf{y}_\sigma^1\}_\sigma$ in $L^\infty(0, T; \mathbf{H}^1(\Omega_h))$, we deduce

$$\begin{aligned} & \frac{1}{2} \|\mathbf{y}_{n,h}\|^2 - \frac{1}{2} \|\mathbf{y}_{n-1,h}\|^2 + \frac{1}{2} \|\mathbf{y}_{n,h} - \mathbf{y}_{n-1,h}\|^2 + \nu \tau_n \|\nabla \mathbf{y}_{n,h}\|^2 \\ &= -\tau_n c(\mathbf{y}_{n,h}, \mathbf{y}_{n,h}^1, \mathbf{y}_{n,h}) \leq \tau_n \|\nabla \mathbf{y}_{n,h}^1\| \|\mathbf{y}_{n,h}\| \|\nabla \mathbf{y}_{n,h}\| \\ &\leq \frac{\nu \tau_n}{2} \|\nabla \mathbf{y}_{n,h}\|^2 + \frac{1}{2\nu} \|\nabla \mathbf{y}_{n,h}^1\|^2 \tau_n \|\mathbf{y}_{n,h}\|^2, \end{aligned}$$

and hence

$$(1 - C\tau_n) \|\mathbf{y}_{n,h}\|^2 + \|\mathbf{y}_{n,h} - \mathbf{y}_{n-1,h}\|^2 + \nu \tau_n \|\nabla \mathbf{y}_{n,h}\|^2 \leq \|\mathbf{y}_{n-1,h}\|^2.$$

Using once again the discrete Gronwall inequality and the fact that $\mathbf{y}_{0,h} = 0$, we conclude that $\mathbf{y}_\sigma = 0$. \square

Remark 4.8. The estimates (4.16) and (4.18) cannot be improved within our optimal control setting. This is due to the regularity restrictions imposed by the nature of problem. However, if $\mathbf{y}_t \in L^2[0, T; H^1(\Omega)]$, then the assumption $\tau \leq Ch^2$ can be dropped (see, e.g., [9]) and the estimate read as $\mathcal{O}(\tau + h)$. However, it is expected that improved estimates in the $L^2[0, T; L^2(\Omega)]$ -norms hold, using an appropriate duality argument. We will examine this issue in a subsequent work. Finally, we remark that discontinuous time-stepping schemes for linear problems typically exhibit nodal (in time) superconvergence (see, e.g., [32] and the references within), under enhanced regularity assumptions. However, it is not clear whether such properties hold, even for the uncontrolled Navier–Stokes equations, with smooth solutions.

Hereafter, we will assume

$$(4.21) \quad \exists C_0 > 0 \text{ such that } \tau \leq C_0 h^2 \quad \forall \sigma = (\tau, h).$$

We establish a corollary of Theorem 4.7 that will be useful later.

COROLLARY 4.9. *Assume that $\max\{\|\mathbf{u}\|_{L^2(0,T;\mathbf{L}^2(\Omega))}, \|\mathbf{v}\|_{L^2(0,T;\mathbf{L}^2(\Omega))}\} \leq M$. Let $\mathbf{y}_u \in \mathbf{H}^{2,1}(\Omega_T) \cap C([0, T]; \mathbf{Y})$ be the solution of (2.1) and $\mathbf{y}_\sigma(\mathbf{v}) \in \mathcal{Y}_\sigma$ the solution of the discrete equation (4.8) corresponding to the control \mathbf{v} . Then there exists a constant $C_M > 0$ such that*

$$(4.22) \quad \begin{aligned} & \|\mathbf{y}_u - \mathbf{y}_\sigma(\mathbf{v})\|_{L^\infty(0,T;\mathbf{L}^2(\Omega_h))} + \|\mathbf{y}_u - \mathbf{y}_\sigma(\mathbf{v})\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))} \\ & \leq C_M \{h + \|\mathbf{u} - \mathbf{v}\|_{L^2(0,T;\mathbf{L}^2(\Omega))}\}. \end{aligned}$$

Moreover, if $\mathbf{u}_\sigma \in \mathcal{U}_\sigma$ for every σ and $\mathbf{u}_\sigma \rightharpoonup \mathbf{u}$ weakly in $L^2(0, T; \mathbf{L}^2(\Omega))$, then

$$(4.23) \quad \begin{cases} \|\mathbf{y}_u - \mathbf{y}_\sigma(\mathbf{u}_\sigma)\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))} \rightarrow 0, \\ \|\mathbf{y}_u - \mathbf{y}_\sigma(\mathbf{u}_\sigma)\|_{L^p(0,T;\mathbf{L}^2(\Omega_h))} \rightarrow 0 \quad \forall 1 \leq p < +\infty, \\ \|\mathbf{y}_u(T) - \mathbf{y}_\sigma(\mathbf{u}_\sigma)(T)\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))} \rightarrow 0. \end{cases}$$

Proof. From (4.18) and (4.21), we get

$$\begin{aligned} \|\mathbf{y}_u - \mathbf{y}_\sigma(\mathbf{v})\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))} &\leq \|\mathbf{y}_u - \mathbf{y}_v\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))} + \|\mathbf{y}_v - \mathbf{y}_\sigma(\mathbf{v})\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))} \\ &\leq \|G(\mathbf{u}) - G(\mathbf{v})\|_{L^2(0,T;\mathbf{H}^1(\Omega))} + Ch, \end{aligned}$$

where C depends on $\|\mathbf{y}_\Omega\|_{\mathbf{H}^1(\Omega)}$ and $\|\mathbf{y}_\mathbf{v}\|_{\mathbf{H}^{2,1}(\Omega_T)}$, bounded by $\|\mathbf{v}\|_{L^2(0,T;\mathbf{L}^2(\Omega))}$. On the other hand, since $G : L^2(0,T;\mathbf{L}^2(\Omega)) \rightarrow \mathbf{H}^{2,1}(\Omega_T) \cap C([0,T];\mathbf{Y})$ is of class C^∞ , we can apply the mean value theorem to get (4.22), with C_M depending on M . Using (4.19), we can repeat the same argument to get the estimate in $L^\infty(0,T;\mathbf{L}^2(\Omega_h))$.

To prove (4.23) we set $\mathbf{y}_\mathbf{u} - \mathbf{y}_\sigma(\mathbf{u}_\sigma) = (\mathbf{y}_\mathbf{u} - \mathbf{y}_{\mathbf{u}_\sigma}) + (\mathbf{y}_{\mathbf{u}_\sigma} - \mathbf{y}_\sigma(\mathbf{u}_\sigma))$. From the well-known properties of (2.1) and the boundedness of $\{\mathbf{f} + \mathbf{u}_\sigma\}_\sigma$ in $L^2(0,T;\mathbf{L}^2(\Omega))$, we have that $\|\mathbf{y}_{\mathbf{u}_\sigma}\|_{\mathbf{H}^{2,1}(\Omega_T)} \leq C$. Furthermore, any subsequence of $\{\mathbf{y}_{\mathbf{u}_\sigma}\}_\sigma$ weakly convergent in $\mathbf{H}^{2,1}(\Omega_T)$ converges to $\mathbf{y}_\mathbf{u}$. This is easily proved by passing to the limit in (2.1). Then we have that $\mathbf{y}_{\mathbf{u}_\sigma} \rightharpoonup \mathbf{y}_\mathbf{u}$ weakly in $\mathbf{H}^{2,1}(\Omega_T)$. From the compactness of the embeddings $\mathbf{H}^{2,1}(\Omega_T) \subset L^2(0,T;\mathbf{H}^1(\Omega))$ and $\mathbf{H}^{2,1}(\Omega_T) \subset L^p(0,T;\mathbf{L}^2(\Omega))$ ($1 \leq p < +\infty$) and the compactness of the trace $\mathbf{H}^{2,1}(\Omega_T) \hookrightarrow \mathbf{L}^2(\partial\Omega_T)$ we obtain

$$\|\mathbf{y}_\mathbf{u} - \mathbf{y}_{\mathbf{u}_\sigma}\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))} + \|\mathbf{y}_\mathbf{u} - \mathbf{y}_{\mathbf{u}_\sigma}\|_{L^p(0,T;\mathbf{L}^2(\Omega_h))} + \|\mathbf{y}_\mathbf{u}(T) - \mathbf{y}_{\mathbf{u}_\sigma}(T)\|_{\mathbf{L}^2(\Omega_h)} \rightarrow 0.$$

On the other hand, from (4.18) and (4.21) we get

$$\|\mathbf{y}_{\mathbf{u}_\sigma}(T) - \mathbf{y}_\sigma(\mathbf{u}_\sigma)(T)\| + \|\mathbf{y}_{\mathbf{u}_\sigma} - \mathbf{y}_\sigma(\mathbf{u}_\sigma)\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))} \rightarrow 0,$$

and with (4.19), $\|\mathbf{y}_{\mathbf{u}_\sigma} - \mathbf{y}_\sigma(\mathbf{u}_\sigma)\|_{L^\infty(0,T;\mathbf{L}^2(\Omega_h))} \rightarrow 0$, which combined with the established convergences imply (4.23). \square

We finish this section studying the differentiability of the relation $\mathbf{u} \rightarrow \mathbf{y}_\sigma(\mathbf{u})$.

THEOREM 4.10. *The mapping $G_\sigma : L^2(0,T;\mathbf{L}^2(\Omega)) \rightarrow \mathcal{Y}_\sigma$, defined by $G_\sigma(\mathbf{u}) = \mathbf{y}_\sigma(\mathbf{u})$ solution of (4.8), is of class C^∞ . Moreover, $\mathbf{z}_\sigma(\mathbf{v}) = G'_\sigma(\mathbf{u})\mathbf{v}$ is the unique solution of the following problem:*

$$(4.24) \quad \begin{cases} \text{For } n = 1, \dots, N_\tau \text{ and } \forall \mathbf{w}_h \in \mathbf{Y}_h, \\ \left(\frac{\mathbf{z}_{n,h} - \mathbf{z}_{n-1,h}}{\tau_n}, \mathbf{w}_h \right) + a(\mathbf{z}_{n,h}, \mathbf{w}_h) + c(\mathbf{z}_{n,h}, \mathbf{y}_{n,h}, \mathbf{w}_h) \\ \quad + c(\mathbf{y}_{n,h}, \mathbf{z}_{n,h}, \mathbf{w}_h) = \frac{1}{\tau_n} \int_{t_{n-1}}^{t_n} (\mathbf{v}(t), \mathbf{w}_h) dt, \\ \mathbf{z}_{0,h} = 0, \end{cases}$$

where we have set $\mathbf{y}_\sigma = \mathbf{y}_\sigma(\mathbf{u})$.

Proof. Let us consider the mapping $F_\sigma : \mathcal{Y}_\sigma \times L^2(0,T;\mathbf{L}^2(\Omega)) \rightarrow \mathcal{Y}'_\sigma$, $F_\sigma(\mathbf{y}_\sigma, \mathbf{u}) = \mathbf{g}_\sigma$, where \mathbf{g}_σ is defined by

$$\begin{aligned} \langle \mathbf{g}_\sigma, \mathbf{w}_\sigma \rangle &= \sum_{n=1}^{N_\tau} \{ (\mathbf{y}_{n,h} - \mathbf{y}_{n-1,h}, \mathbf{w}_{n,h}) + \tau_n [a(\mathbf{y}_{n,h}, \mathbf{w}_{n,h}) + c(\mathbf{y}_{n,h}, \mathbf{y}_{n,h}, \mathbf{w}_{n,h})] \} \\ &\quad - \sum_{n=1}^{N_\tau} \int_{t_{n-1}}^{t_n} (\mathbf{f}(t) + \mathbf{u}(t), \mathbf{w}_{n,h}) dt \quad \forall \mathbf{w}_\sigma \in \mathcal{Y}_\sigma. \end{aligned}$$

Obviously, F_σ is of class C^∞ and $\hat{\mathbf{g}}_\sigma = \frac{\partial F_\sigma}{\partial \mathbf{y}_\sigma}(\mathbf{y}_\sigma, \mathbf{u})\mathbf{z}_\sigma$ is defined by

$$\begin{aligned} \langle \hat{\mathbf{g}}, \mathbf{w}_\sigma \rangle &= \sum_{n=1}^{N_\tau} \{ (\mathbf{z}_{n,h} - \mathbf{z}_{n-1,h}, \mathbf{w}_{n,h}) + \tau_n a(\mathbf{z}_{n,h}, \mathbf{w}_{n,h}) \} \\ &\quad + \sum_{n=1}^{N_\tau} \{ \tau_n [c(\mathbf{y}_{n,h}, \mathbf{z}_{n,h}, \mathbf{w}_{n,h}) + c(\mathbf{z}_{n,h}, \mathbf{y}_{n,h}, \mathbf{w}_{n,h})] \}, \quad \text{with } \mathbf{z}_{0,h} = 0. \end{aligned}$$

On the other hand, $F_\sigma(G_\sigma(\mathbf{u}), \mathbf{u}) = F_\sigma(\mathbf{y}_\sigma(\mathbf{u}), \mathbf{u}) = 0$ for every $\mathbf{u} \in L^2(0,T;\mathbf{L}^2(\Omega))$. The proof is a consequence of the implicit function theorem; we need to prove that

$\frac{\partial F_\sigma}{\partial \mathbf{y}_\sigma}(\mathbf{y}_\sigma(\mathbf{u}), \mathbf{u}) : \mathcal{Y}_\sigma \longrightarrow \mathcal{Y}'_\sigma$ is an isomorphism for every \mathbf{u} . In fact, we will prove that $\frac{\partial F_\sigma}{\partial \mathbf{y}_\sigma}(\mathbf{y}_\sigma, \mathbf{u})$ is an isomorphism for every $(\mathbf{y}_\sigma, \mathbf{u}) \in \mathcal{Y}_\sigma \times L^2(0, T; \mathbf{L}^2(\Omega))$. Since $\frac{\partial F_\sigma}{\partial \mathbf{y}_\sigma}(\mathbf{y}_\sigma, \mathbf{u})$ is a linear mapping between two spaces of the same finite dimension, it is enough to prove that it is injective. Suppose that $\frac{\partial F_\sigma}{\partial \mathbf{y}_\sigma}(\mathbf{y}_\sigma, \mathbf{u})\mathbf{z}_\sigma = 0$ for some $\mathbf{z}_\sigma \in \mathcal{Y}_\sigma$. Applying $\frac{\partial F_\sigma}{\partial \mathbf{y}_\sigma}(\mathbf{y}_\sigma, \mathbf{u})\mathbf{z}_\sigma \in \mathcal{Y}'_\sigma$ to \mathbf{z}_σ and using that $c(\mathbf{y}_{n,h}, \mathbf{z}_{n,h}, \mathbf{z}_{n,h}) = 0$, we get

$$\sum_{n=1}^{N_\tau} \{(\mathbf{z}_{n,h} - \mathbf{z}_{n-1,h}, \mathbf{z}_{n,h}) + \tau_n [a(\mathbf{z}_{n,h}, \mathbf{z}_{n,h}) + c(\mathbf{z}_{n,h}, \mathbf{y}_{n,h}, \mathbf{z}_{n,h})]\} = 0.$$

Hence,

$$\begin{aligned} & \sum_{n=1}^{N_\tau} \left\{ \frac{1}{2} \|\mathbf{z}_{n,h}\|^2 - \frac{1}{2} \|\mathbf{z}_{n-1,h}\|^2 + \frac{1}{2} \|\mathbf{z}_{n,h} - \mathbf{z}_{n-1,h}\|^2 + \nu \tau_n \|\nabla \mathbf{z}_{n,h}\|^2 \right\} \\ & \leq \sum_{n=1}^{N_\tau} \tau_n \|\nabla \mathbf{y}_{n,h}\| \|\mathbf{z}_{n,h}\| \|\nabla \mathbf{z}_{n,h}\| \\ & \leq \frac{1}{\nu} \|\mathbf{y}_\sigma\|_{L^\infty(0,T;\mathbf{H}^1(\Omega_h))} \sum_{n=1}^{N_\tau} \tau_n \|\mathbf{z}_{n,h}\|^2 + \frac{\nu}{2} \sum_{n=1}^{N_\tau} \tau_n \|\nabla \mathbf{z}_{n,h}\|^2. \end{aligned}$$

Again, an application of the discrete Gronwall inequality and the fact that $\mathbf{z}_{0,h} = 0$ imply that $\mathbf{z}_\sigma = 0$. \square

4.2. Analysis of the discrete adjoint state equation. In this section, as well as in the rest of the paper, the condition (4.21) is assumed. As a consequence of Theorem 4.10 and applying the chain rule, we get that $J_\sigma : L^2(0, T; \mathbf{L}^2(\Omega)) \longrightarrow \mathbb{R}$ is of class C^∞ , and we have a first expression of its derivative as follows:

$$\begin{aligned} J'_\sigma(\mathbf{u})\mathbf{v} &= \int_0^T \int_{\Omega_h} (\mathbf{y}_\sigma - \mathbf{y}_d) \mathbf{z}_\sigma \, dx \, dt \\ &+ \gamma \int_{\Omega_h} (\mathbf{y}_\sigma(T) - \mathbf{y}_d) \mathbf{z}_\sigma(T) \, dx + \lambda \int_0^T \int_{\Omega_h} \mathbf{u} \mathbf{v} \, dx \, dt, \end{aligned}$$

where $\mathbf{y}_\sigma = \mathbf{y}_\sigma(\mathbf{u}) = G_\sigma(\mathbf{u})$ and $\mathbf{z}_\sigma = G'_\sigma(\mathbf{u})\mathbf{v}$ is the solution of (4.24). As usual in control theory, we have to introduce the adjoint state to simplify the expression of this derivative. To this end we consider the discrete adjoint state equation: We look for $\varphi_\sigma \in \mathcal{Y}_\sigma$ such that

$$(4.25) \quad \begin{cases} \text{for } n = N_\tau, \dots, 1 \text{ and } \forall \mathbf{w}_h \in \mathbf{Y}_h, \\ \left(\frac{\varphi_{n,h} - \varphi_{n+1,h}}{\tau_n}, \mathbf{w}_h \right) + a(\varphi_{n,h}, \mathbf{w}_h) + c(\mathbf{w}_h, \mathbf{y}_{n,h}, \varphi_{n,h}) \\ + c(\mathbf{y}_{n,h}, \mathbf{w}_h, \varphi_{n,h}) = \frac{1}{\tau_n} \int_{t_{n-1}}^{t_n} (\mathbf{y}_{n,h} - \mathbf{y}_d(t), \mathbf{w}_h) \, dt, \\ \varphi_{N_\tau+1,h} = \gamma(\mathbf{y}_{N_\tau,h} - \mathbf{y}_{\Omega_h}). \end{cases}$$

Observe that in the above system, first we compute $\varphi_{N_\tau,h}$ from $\varphi_{N_\tau+1,h} = \gamma(\mathbf{y}_{N_\tau,h} - \mathbf{y}_{\Omega_h})$ and then we descend in n until $n = 1$. Unlike the discrete states \mathbf{y}_σ , we will set for the discrete adjoint states $\varphi_\sigma(t_{n-1}) = \varphi_{n,h}$ for every $1 \leq n \leq N_\tau$.

System (4.25) corresponds to the discretization of the backward equation (3.5). Using that $\{\mathbf{y}_\sigma\}_\sigma$ is bounded in $L^\infty(0, T; \mathbf{H}^1(\Omega_h))$ (Theorem (4.7)), then we can

proceed in the same way as we did in the proof of Theorem 4.10 to obtain the existence and uniqueness of a solution of (4.25). Below we check that this is actually the discrete adjoint state equation. To this end we use (4.24) and (4.25) to show that

$$\begin{aligned}
 & \int_0^T \int_{\Omega_h} (\mathbf{y}_\sigma - \mathbf{y}_d) \mathbf{z}_\sigma \, dx \, dt \\
 &= \sum_{n=1}^{N_\tau} \int_{t_{n-1}}^{t_n} (\mathbf{y}_{n,h} - \mathbf{y}_d(t), \mathbf{z}_{n,h}) \, dt = \sum_{n=1}^{N_\tau} (\varphi_{n,h} - \varphi_{n+1,h}, \mathbf{z}_{n,h}) \\
 &+ \sum_{n=1}^{N_\tau} \tau_n [a(\varphi_{n,h}, \mathbf{z}_{n,h}) + c(\mathbf{z}_{n,h}, \mathbf{y}_{n,h}, \varphi_{n,h}) + c(\mathbf{y}_{n,h}, \mathbf{z}_{n,h}, \varphi_{n,h})] \\
 &= \sum_{n=1}^{N_\tau} (\mathbf{z}_{n,h} - \mathbf{z}_{n-1,h}, \varphi_{n,h}) - (\varphi_{N_\tau+1,h}, \mathbf{z}_{N_\tau,h}) + (\varphi_{1,h}, \mathbf{z}_{0,h}) \\
 &+ \sum_{n=1}^{N_\tau} \tau_n [a(\mathbf{z}_{n,h}, \varphi_{n,h}) + c(\mathbf{z}_{n,h}, \mathbf{y}_{n,h}, \varphi_{n,h}) + c(\mathbf{y}_{n,h}, \mathbf{z}_{n,h}, \varphi_{n,h})] \\
 &= \int_0^T \int_{\Omega_h} \mathbf{v} \varphi_\sigma \, dx \, dt - \gamma \int_{\Omega_h} (\mathbf{y}_\sigma(T) - \mathbf{y}_{\Omega_h}) \mathbf{z}_\sigma(T) \, dx,
 \end{aligned}$$

where we have used that $\varphi_{N_\tau+1,h} = \gamma(\mathbf{y}_{N_\tau,h} - \mathbf{y}_{\Omega_h}) = \gamma(\mathbf{y}_\sigma(T) - \mathbf{y}_{\Omega_h})$ and $\mathbf{z}_{0,h} = 0$. From the obtained identity and the expression of $J'_\sigma(\mathbf{u})\mathbf{v}$ given above we conclude

$$(4.26) \quad J'_\sigma(\mathbf{u})\mathbf{v} = \int_0^T \int_{\Omega_h} (\varphi_\sigma + \lambda \mathbf{u}) \mathbf{v} \, dx \, dt.$$

The next theorem states the error estimates in the approximation of the adjoint state equation.

THEOREM 4.11. *Given $\mathbf{u} \in L^2(0, T; \mathbf{L}^2(\Omega))$, let $\mathbf{y} = \mathbf{y}_\mathbf{u}$ be the associated state, solution of (2.1), φ the associated adjoint state, solution of (3.5), $\mathbf{y}_\sigma = \mathbf{y}_\sigma(\mathbf{u})$ the associated discrete state, solution of (4.8), and φ_σ the associated discrete adjoint state, solution of (4.25). Then $\{\varphi_\sigma\}_\sigma$ is bounded in $L^\infty(0, T; \mathbf{H}^1(\Omega_h))$ and there exists a constant $C > 0$ independent of σ and \mathbf{u} such that*

$$(4.27) \quad \begin{aligned} & \|\varphi - \varphi_\sigma\|_{L^\infty(0, T; \mathbf{L}^2(\Omega_h))} + \|\varphi - \varphi_\sigma\|_{L^2(0, T; \mathbf{H}^1(\Omega_h))} \\ & \leq Ch \left\{ \|\mathbf{u}\|_{L^2(0, T; \mathbf{L}^2(\Omega))} + \|\mathbf{y}_0\|_{\mathbf{H}^1(\Omega)} + \|\mathbf{y}_\Omega\|_{\mathbf{H}^1(\Omega)} \right\}. \end{aligned}$$

Proof. Define the operator $R_\sigma : C([0, T]; \mathbf{L}^2(\Omega)) \rightarrow \mathcal{Y}_\sigma$ by $(R_\sigma \mathbf{w})_{n,h} = P_h \mathbf{w}(t_{n-1})$ for $1 \leq n \leq N_\tau$, with P_h given in Definition 4.1. As for the discrete adjoint states, we fix $(R_\sigma \mathbf{w})(t_{n-1}) = (R_\sigma \mathbf{w})_{n,h}$. Analogously to (4.12) and (4.14), we have the following estimates for every $\mathbf{w} \in \mathbf{H}^{2,1}(\Omega_T) \cap C([0, T], \mathbf{Y})$:

$$(4.28) \quad \|\mathbf{w} - R_\sigma \mathbf{w}\|_{L^2(0, T; \mathbf{L}^2(\Omega_h))} \leq C \left\{ \tau \|\mathbf{w}'\|_{L^2(0, T; \mathbf{L}^2(\Omega))} + h^2 \|\mathbf{w}\|_{L^2(0, T; \mathbf{H}^2(\Omega))} \right\},$$

$$(4.29) \quad \|\mathbf{w} - R_\sigma \mathbf{w}\|_{L^2(0, T; \mathbf{H}^1(\Omega_h))} \leq C \left\{ \frac{\tau}{h} \|\mathbf{w}'\|_{L^2(0, T; \mathbf{L}^2(\Omega))} + h \|\mathbf{w}\|_{L^2(0, T; \mathbf{H}^2(\Omega))} \right\}.$$

We set $\epsilon = \varphi - \varphi_\sigma = (\varphi - R_\sigma \varphi) + (R_\sigma \varphi - \varphi_\sigma) = \boldsymbol{\eta} + \boldsymbol{\epsilon}_\sigma$. According to our fixed notation above, we have $\boldsymbol{\eta}(t_n) = \varphi(t_n) - (R_\sigma \varphi)(t_n) = \varphi(t_n) - (R_\sigma \varphi)_{n+1,h} = \varphi(t_n) - P_h \varphi(t_n)$ for $0 \leq n \leq N_\tau - 1$. Also we have $\boldsymbol{\epsilon}_\sigma(t_n) = \boldsymbol{\epsilon}_{n+1,h}$, $0 \leq n \leq N_\tau - 1$. Setting $(R_\sigma \mathbf{w})_{N_\tau+1,h} = P_h \mathbf{w}(T)$ and recalling that $\varphi_{N_\tau+1,h} = \gamma(\mathbf{y}_{N_\tau,h} - \mathbf{y}_{\Omega_h})$, the previous

identities are also well defined for $n = N_\tau$. Then (3.5) and (4.25) lead to the identities, $n = N_\tau, \dots, 1$,

$$\begin{aligned} & (\epsilon(t_{n-1}) - \epsilon(t_n), \mathbf{w}_h) + \int_{t_{n-1}}^{t_n} a(\epsilon(t), \mathbf{w}_h) dt \\ & + \int_{t_{n-1}}^{t_n} [c(\mathbf{y}(t), \mathbf{w}_h, \varphi(t)) + c(\mathbf{w}_h, \mathbf{y}(t), \varphi(t))] dt \\ & - \int_{t_{n-1}}^{t_n} [c(\mathbf{y}_{n,h}, \mathbf{w}_h, \varphi_{n,h}) + c(\mathbf{w}_h, \mathbf{y}_{n,h}, \varphi_{n,h})] dt = \int_{t_{n-1}}^{t_n} (\mathbf{y}(t) - \mathbf{y}_{n,h}, \mathbf{w}_h) dt. \end{aligned}$$

Now, writing $\epsilon = \eta + \epsilon_\sigma$ and taking into account that

$$(\eta(t_n), \mathbf{w}_h) = (\varphi(t_n) - P_h \varphi(t_n), \mathbf{w}_h) = 0 \quad \forall \mathbf{w}_h \in \mathbf{Y}_h \text{ and for } 0 \leq n \leq N_\tau,$$

we obtain for $\mathbf{w}_h = \epsilon_{n,h}$

$$\begin{aligned} & (\epsilon_{n,h} - \epsilon_{n+1,h}, \epsilon_{n,h}) + \int_{t_{n-1}}^{t_n} a(\epsilon_{n,h}, \epsilon_{n,h}) dt \\ & = \int_{t_{n-1}}^{t_n} (\mathbf{y}(t) - \mathbf{y}_{n,h}, \epsilon_{n,h}) dt - \int_{t_{n-1}}^{t_n} a(\eta(t), \epsilon_{n,h}) dt \\ & + \int_{t_{n-1}}^{t_n} [c(\mathbf{y}_{n,h}, \epsilon_{n,h}, \varphi_{n,h}) + c(\epsilon_{n,h}, \mathbf{y}_{n,h}, \varphi_{n,h})] dt \\ & - \int_{t_{n-1}}^{t_n} [c(\mathbf{y}(t), \epsilon_{n,h}, \varphi(t)) + c(\epsilon_{n,h}, \mathbf{y}(t), \varphi(t))] dt; \end{aligned}$$

hence

$$\begin{aligned} & \frac{1}{2} \|\epsilon_{n,h}\|^2 - \frac{1}{2} \|\epsilon_{n+1,h}\|^2 + \frac{1}{2} \|\epsilon_{n,h} - \epsilon_{n+1,h}\|^2 + \nu \int_{t_{n-1}}^{t_n} \|\nabla \epsilon_{n,h}\|^2 dt \\ & \leq \int_{t_{n-1}}^{t_n} \|\mathbf{y}(t) - \mathbf{y}_\sigma(t)\| \|\epsilon_{n,h}\| dt + \nu \int_{t_{n-1}}^{t_n} \|\nabla \eta(t)\| \|\nabla \epsilon_{n,h}\| dt \\ & + \int_{t_{n-1}}^{t_n} [c(\mathbf{y}_{n,h}, \epsilon_{n,h}, \varphi_{n,h}) - c(\mathbf{y}(t), \epsilon_{n,h}, \varphi(t))] dt \\ (4.30) \quad & + \int_{t_{n-1}}^{t_n} [c(\epsilon_{n,h}, \mathbf{y}_{n,h}, \varphi_{n,h}) - c(\epsilon_{n,h}, \mathbf{y}(t), \varphi(t))] dt. \end{aligned}$$

Let us estimate the right-hand side of (4.30):

$$\begin{aligned} & \int_{t_{n-1}}^{t_n} \|\mathbf{y}(t) - \mathbf{y}_\sigma(t)\| \|\epsilon_{n,h}\| dt + \nu \int_{t_{n-1}}^{t_n} \|\nabla \eta(t)\| \|\nabla \epsilon_{n,h}\| dt \\ & \leq \frac{1}{2} \int_{t_{n-1}}^{t_n} \|\mathbf{y}(t) - \mathbf{y}_\sigma(t)\|^2 dt + \frac{\tau_n}{2} \|\epsilon_{n,h}\|^2 \\ & + 2\nu \int_{t_{n-1}}^{t_n} \|\nabla \eta(t)\|^2 dt + \frac{\nu}{8} \int_{t_{n-1}}^{t_n} \|\nabla \epsilon_{n,h}\|^2 dt. \end{aligned}$$

Now we proceed with the second term:

$$\begin{aligned}
& \left| \int_{t_{n-1}}^{t_n} [c(\mathbf{y}_{n,h}, \boldsymbol{\epsilon}_{n,h}, \boldsymbol{\varphi}_{n,h}) - c(\mathbf{y}(t), \boldsymbol{\epsilon}_{n,h}, \boldsymbol{\varphi}(t))] dt \right| \\
&= \left| \int_{t_{n-1}}^{t_n} [c(\mathbf{y}_\sigma(t) - \mathbf{y}(t), \boldsymbol{\epsilon}_{n,h}, \boldsymbol{\varphi}(t)) - c(\mathbf{y}_\sigma(t), \boldsymbol{\epsilon}_{n,h}, \boldsymbol{\epsilon}(t))] dt \right| \\
&= \left| \int_{t_{n-1}}^{t_n} [c(\mathbf{y}_\sigma(t) - \mathbf{y}(t), \boldsymbol{\epsilon}_{n,h}, \boldsymbol{\varphi}(t)) - c(\mathbf{y}_\sigma(t), \boldsymbol{\epsilon}_{n,h}, \boldsymbol{\eta}(t))] dt \right| \\
&\leq \|\boldsymbol{\varphi}\|_{L^\infty(0,T;\mathbf{H}^1(\Omega))} \int_{t_{n-1}}^{t_n} \|\mathbf{y}(t) - \mathbf{y}_\sigma(t)\|_{\mathbf{H}^1(\Omega_h)} \|\nabla \boldsymbol{\epsilon}_{n,h}\| dt \\
&\quad + \|\mathbf{y}_\sigma\|_{L^\infty(0,T;\mathbf{H}^1(\Omega_h))} \int_{t_{n-1}}^{t_n} \|\boldsymbol{\eta}(t)\|_{\mathbf{H}^1(\Omega_h)} \|\nabla \boldsymbol{\epsilon}_{n,h}\| dt \\
&\leq C \int_{t_{n-1}}^{t_n} \left[\|\mathbf{y}(t) - \mathbf{y}_\sigma(t)\|_{\mathbf{H}^1(\Omega_h)}^2 + \|\boldsymbol{\eta}(t)\|_{\mathbf{H}^1(\Omega_h)}^2 \right] dt + \frac{\nu}{4} \int_{t_{n-1}}^{t_n} \|\nabla \boldsymbol{\epsilon}_{n,h}\|^2 dt.
\end{aligned}$$

For the last term of (4.30), we first observe that

$$\begin{aligned}
& c(\boldsymbol{\epsilon}_{n,h}, \mathbf{y}_{n,h}, \boldsymbol{\varphi}_{n,h}) - c(\boldsymbol{\epsilon}_{n,h}, \mathbf{y}(t), \boldsymbol{\varphi}(t)) \\
&= -[c(\boldsymbol{\epsilon}_{n,h}, \mathbf{y}(t) - \mathbf{y}_\sigma(t), \boldsymbol{\varphi}(t)) + c(\boldsymbol{\epsilon}_{n,h}, \mathbf{y}_\sigma(t), \boldsymbol{\eta}(t)) + c(\boldsymbol{\epsilon}_{n,h}, \mathbf{y}_\sigma(t), \boldsymbol{\epsilon}_{n,h})].
\end{aligned}$$

The first two terms can be estimated in a similar way to the previous one. For the last we get

$$\begin{aligned}
& \left| \int_{t_{n-1}}^{t_n} c(\boldsymbol{\epsilon}_{n,h}, \mathbf{y}_\sigma(t), \boldsymbol{\epsilon}_{n,h}) dt \right| \leq \sqrt{2} \|\mathbf{y}_\sigma\|_{L^\infty(0,T;\mathbf{H}^1(\Omega_h))} \int_{t_{n-1}}^{t_n} \|\boldsymbol{\epsilon}_{n,h}\| \|\nabla \boldsymbol{\epsilon}_{n,h}\| dt \\
&\leq C \tau_n \|\boldsymbol{\epsilon}_{n,h}\|^2 + \frac{\nu}{8} \int_{t_{n-1}}^{t_n} \|\nabla \boldsymbol{\epsilon}_{n,h}\|^2 dt.
\end{aligned}$$

Collecting all the estimates, we infer from (4.30) that

$$\begin{aligned}
& (1 - C\tau_n) \|\boldsymbol{\epsilon}_{n,h}\|^2 + \|\boldsymbol{\epsilon}_{n,h} - \boldsymbol{\epsilon}_{n+1,h}\|^2 + \frac{\nu}{4} \int_{t_{n-1}}^{t_n} \|\nabla \boldsymbol{\epsilon}_{n,h}\|^2 dt \\
&\leq \|\boldsymbol{\epsilon}_{n+1,h}\|^2 + C \left\{ \int_{t_{n-1}}^{t_n} \|\boldsymbol{\eta}(t)\|_{\mathbf{H}^1(\Omega_h)}^2 dt + \int_{t_{n-1}}^{t_n} \|\mathbf{y}(t) - \mathbf{y}_\sigma(t)\|^2 dt \right\}.
\end{aligned}$$

To conclude the proof it is enough to use the discrete Gronwall inequality along with (4.11), (4.18), (4.21), (4.29), and the fact that the $\mathbf{H}^{2,1}(\Omega_T)$ -norm of $\boldsymbol{\varphi}$ can be estimated by the $L^2(0,T;\mathbf{L}^2(\Omega))$ -norm of $\mathbf{y} - \mathbf{y}_d$ and the $\mathbf{H}^1(\Omega)$ -norm of \mathbf{y}_Ω , and the $L^2(0,T;\mathbf{L}^2(\Omega))$ -norm of \mathbf{y} is estimated by the $L^2(0,T;\mathbf{L}^2(\Omega))$ -norm of \mathbf{u} . \square

As a consequence of the previous theorem we have the following result analogous to Corollary 4.9.

COROLLARY 4.12. Assume that $\max\{\|\mathbf{u}\|_{L^2(0,T;\mathbf{L}^2(\Omega))}, \|\mathbf{v}\|_{L^2(0,T;\mathbf{L}^2(\Omega))}\} \leq M$. Let $\boldsymbol{\varphi}_u \in \mathbf{H}^{2,1}(\Omega_T) \cap C([0,T];\mathbf{Y})$ be the solution of (3.5) and $\boldsymbol{\varphi}_\sigma(\mathbf{v}) \in \mathcal{Y}_\sigma$ the solution of the discrete equation (4.25) corresponding to the control \mathbf{v} . Then there exists a constant $C_M > 0$ such that

$$\begin{aligned}
& \|\boldsymbol{\varphi}_u - \boldsymbol{\varphi}_\sigma(\mathbf{v})\|_{L^\infty(0,T;\mathbf{L}^2(\Omega_h))} + \|\boldsymbol{\varphi}_u - \boldsymbol{\varphi}_\sigma(\mathbf{v})\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))} \\
(4.31) \quad & \leq C_M \{h + \|\mathbf{u} - \mathbf{v}\|_{L^2(0,T;\mathbf{L}^2(\Omega))}\}.
\end{aligned}$$

Proof. First we observe that (4.27) implies

$$\|\varphi_{\mathbf{u}} - \varphi_{\sigma}(\mathbf{v})\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))} \leq \|\varphi_{\mathbf{u}} - \varphi_{\mathbf{v}}\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))} + Ch,$$

where C depends on $\|\mathbf{u}\|_{L^2(0,T;\mathbf{L}^2(\Omega))}$. We proceed analogously to get the estimate for $\|\varphi_{\mathbf{u}} - \varphi_{\sigma}(\mathbf{v})\|_{L^\infty(0,T;\mathbf{L}^2(\Omega_h))}$. Now, we estimate $\varphi_{\mathbf{u}} - \varphi_{\mathbf{v}}$ in $L^2(0,T;\mathbf{H}^1(\Omega))$ and $L^\infty(0,T;\mathbf{L}^2(\Omega))$, respectively. Let us set $\varphi = \varphi_{\mathbf{u}} - \varphi_{\mathbf{v}}$; then subtracting the equations satisfied by $\varphi_{\mathbf{u}}$ and $\varphi_{\mathbf{v}}$, we get

$$\begin{aligned} -(\varphi_t, \mathbf{w}) + a(\varphi, \mathbf{w}) &= (\mathbf{y}_{\mathbf{u}} - \mathbf{y}_{\mathbf{v}}, \mathbf{w}) \\ &+ c(\mathbf{w}, \mathbf{y}_{\mathbf{v}}, \varphi_{\mathbf{v}}) + c(\mathbf{y}_{\mathbf{v}}, \mathbf{w}, \varphi_{\mathbf{v}}) - c(\mathbf{w}, \mathbf{y}_{\mathbf{u}}, \varphi_{\mathbf{u}}) - c(\mathbf{y}_{\mathbf{u}}, \mathbf{w}, \varphi_{\mathbf{u}}). \end{aligned}$$

Taking $\mathbf{w} = \varphi$ and using the identities

$$\begin{aligned} c(\varphi, \mathbf{y}_{\mathbf{v}}, \varphi_{\mathbf{v}}) - c(\varphi, \mathbf{y}_{\mathbf{u}}, \varphi_{\mathbf{u}}) &= c(\varphi, \mathbf{y}_{\mathbf{v}} - \mathbf{y}_{\mathbf{u}}, \varphi_{\mathbf{v}}) - c(\varphi, \mathbf{y}_{\mathbf{u}}, \varphi), \\ c(\mathbf{y}_{\mathbf{v}}, \varphi, \varphi_{\mathbf{v}}) - c(\mathbf{y}_{\mathbf{u}}, \varphi, \varphi_{\mathbf{u}}) &= c(\mathbf{y}_{\mathbf{v}} - \mathbf{y}_{\mathbf{u}}, \varphi, \varphi_{\mathbf{v}}), \end{aligned}$$

we deduce by integration in the interval (t, T) and the equality $\varphi(T) = \varphi_{\mathbf{u}}(T) - \varphi_{\mathbf{v}}(T) = \gamma(\mathbf{y}_{\mathbf{u}}(T) - \mathbf{y}_{\mathbf{v}}(T))$ that

$$\begin{aligned} &\frac{1}{2}\|\varphi(t)\|^2 - \gamma^2\frac{1}{2}\|\mathbf{y}_{\mathbf{u}}(T) - \mathbf{y}_{\mathbf{v}}(T)\|^2 + \nu \int_t^T \|\nabla \varphi(s)\|^2 ds \\ &\leq \int_t^T \|\mathbf{y}_{\mathbf{u}}(s) - \mathbf{y}_{\mathbf{v}}(s)\| \|\varphi(s)\| ds \\ &+ C \int_t^T \left\{ \|\varphi(s)\|^{1/2} \|\nabla \varphi(s)\|^{1/2} \|\nabla \mathbf{y}_{\mathbf{v}}(s) - \nabla \mathbf{y}_{\mathbf{u}}(s)\| \|\varphi_{\mathbf{v}}(s)\|_{\mathbf{H}^1(\Omega)} \right\} ds \\ &+ C \int_t^T \|\varphi(s)\| \|\nabla \varphi(s)\| \|\nabla \mathbf{y}_{\mathbf{u}}\| ds \\ &+ C \int_t^T \|\mathbf{y}_{\mathbf{v}}(s) - \mathbf{y}_{\mathbf{u}}(s)\|_{\mathbf{H}^1(\Omega)} \|\nabla \varphi(s)\| \|\varphi_{\mathbf{v}}(s)\|_{\mathbf{H}^1(\Omega)} ds. \end{aligned}$$

Since $\mathbf{y}_{\mathbf{u}}, \varphi_{\mathbf{v}} \in L^\infty(0, T; \mathbf{H}^1(\Omega))$, with norms estimated by a constant depending on M , we infer from the above inequality that

$$\begin{aligned} &\frac{1}{2}\|\varphi(t)\|^2 - \gamma^2\frac{1}{2}\|\mathbf{y}_{\mathbf{u}}(T) - \mathbf{y}_{\mathbf{v}}(T)\|^2 + \nu \int_t^T \|\nabla \varphi(s)\|^2 ds \\ &\leq \frac{1}{2} \int_t^T \|\mathbf{y}_{\mathbf{u}}(s) - \mathbf{y}_{\mathbf{v}}(s)\|^2 ds + \frac{1}{2} \int_t^T \|\varphi(s)\|^2 ds \\ &+ C \int_t^T \left\{ \|\varphi(s)\|^2 + \|\mathbf{y}_{\mathbf{v}}(s) - \mathbf{y}_{\mathbf{u}}(s)\|_{\mathbf{H}^1(\Omega)}^2 \right\} ds + \frac{\nu}{2} \int_t^T \|\nabla \varphi(s)\|^2 ds. \end{aligned}$$

On the other hand, we have

$$\begin{aligned} \|\mathbf{y}_{\mathbf{u}} - \mathbf{y}_{\mathbf{v}}\|_{\mathbf{H}^{2,1}(\Omega_T)} &= \|G(\mathbf{u}) - G(\mathbf{v})\|_{\mathbf{H}^{2,1}(\Omega_T)} \\ &\leq \sup_{0 \leq \rho \leq 1} \|G'(\mathbf{u} + \rho(\mathbf{v} - \mathbf{u}))\| \|\mathbf{u} - \mathbf{v}\|_{L^2(0,T;\mathbf{L}^2(\Omega))} \leq C \|\mathbf{u} - \mathbf{v}\|_{L^2(0,T;\mathbf{L}^2(\Omega))}, \end{aligned}$$

where C depends on M . The last two inequalities lead to

$$\begin{aligned} & \|\varphi(t)\|^2 + \nu \int_t^T \|\nabla \varphi(s)\|^2 ds \\ & \leq C \left\{ \|\mathbf{y}_u(T) - \mathbf{y}_v(T)\|^2 + \|\mathbf{y}_u - \mathbf{y}_v\|_{L^2(0,T;\mathbf{H}^1(\Omega))}^2 + \int_t^T \|\varphi(s)\|^2 ds \right\} \\ & \leq C \left\{ \|\mathbf{y}_u - \mathbf{y}_v\|_{\mathbf{H}^{2,1}(\Omega_T)}^2 + \int_t^T \|\varphi(s)\|^2 ds \right\} \\ & \leq C \left\{ \|\mathbf{u} - \mathbf{v}\|_{L^2(0,T;\mathbf{L}^2(\Omega))}^2 + \int_t^T \|\varphi(s)\|^2 ds \right\} \quad \forall t \in [0, T]. \end{aligned}$$

Now the Gronwall inequality implies

$$\|\varphi(t)\| \leq C \|\mathbf{u} - \mathbf{v}\|_{L^2(0,T;\mathbf{L}^2(\Omega))} \quad \forall t \in [0, T],$$

which also implies with the aid of the previous estimates that

$$\nu \int_0^T \|\nabla \varphi(s)\|^2 ds \leq C \|\mathbf{u} - \mathbf{v}\|_{L^2(0,T;\mathbf{L}^2(\Omega))}^2,$$

which concludes the proof. \square

4.3. Convergence of the discrete control problem. In this section we analyze the convergence of the solutions of control problems (P_σ) towards solutions of the continuous problem (P) . Since these problems are not convex, we will also address the issue of the approximation of local solutions of problem (P) . It is clear that every problem (P_σ) has at least one solution because it consists of the minimization of a continuous and coercive function on a nonempty closed subset of a finite dimensional space. The next theorem proves the convergence of these discrete solutions to solutions of problem (P) .

THEOREM 4.13. *For every $\sigma = (\tau, h)$ let $\bar{\mathbf{u}}_\sigma$ be a global solution of problem (P_σ) . Then the sequence $\{\bar{\mathbf{u}}_\sigma\}_\sigma$ is bounded in $L^2(0, T; \mathbf{L}^2(\Omega))$ and there exist subsequences, denoted in the same way, converging to a point $\bar{\mathbf{u}}$ weakly in $L^2(0, T; \mathbf{L}^2(\Omega))$. Any of these limit points is a solution of problem (P) . Moreover, we have*

$$(4.32) \quad \lim_{\sigma \rightarrow 0} \|\bar{\mathbf{u}} - \bar{\mathbf{u}}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))} = 0 \quad \text{and} \quad \lim_{\sigma \rightarrow 0} J_\sigma(\bar{\mathbf{u}}_\sigma) = J(\bar{\mathbf{u}}).$$

Remark 4.14. Strictly speaking, it is not correct to claim that sequence $\{\bar{\mathbf{u}}_\sigma\}_\sigma$ is bounded in $L^2(0, T; \mathbf{L}^2(\Omega))$ because $\bar{\mathbf{u}}_\sigma$ is only defined in $(0, T) \times \Omega_h$, with $\Omega_h \subset \Omega$, for $\sigma = (\tau, h)$. We will prove that $\|\mathbf{u}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))} \leq C$ for some constant independent of σ . Now, if we take any element $\mathbf{v} \in L^2(0, T; \mathbf{L}^2(\Omega))$ and extend every $\bar{\mathbf{u}}_\sigma$ to $(0, T) \times \Omega$ by setting $\bar{\mathbf{u}}_\sigma(t, x) = \mathbf{v}(t, x)$ for every $(t, x) \in (0, T) \times (\Omega \setminus \Omega_h)$, then we have that these extensions constitute a sequence of bounded functions in $L^2(0, T; \mathbf{L}^2(\Omega))$ and every weak limit point is a solution of (P) , regardless of the choice of \mathbf{v} . This is a consequence of the property (4.1). The theorem should be understood in this sense.

Proof. Let $\bar{\mathbf{u}}$ be a solution of problem (P) and let us take $\mathbf{u}_\sigma \in \mathcal{U}_\sigma$ defined by

$$(4.33) \quad \mathbf{u}_\sigma = \sum_{n=1}^{N_\tau} \sum_{T \in \mathcal{T}_h} \mathbf{u}_{n,T} \chi_n \chi_T, \quad \text{with} \quad \mathbf{u}_{n,T} = \frac{1}{\tau_n |T|} \int_{t_{n-1}}^{t_n} \int_T \tilde{\mathbf{u}}(t, x) dx dt.$$

Then \mathbf{u}_σ is the $L^2(0, T; \mathbf{L}^2(\Omega_h))$ projection of $\tilde{\mathbf{u}}$ on \mathcal{Y}_σ . From our assumptions (A1)–(A3), we have that $\|\tilde{\mathbf{u}} - \mathbf{u}_\sigma\|_{L^2(0, T; \mathbf{L}^2(\Omega_h))} \rightarrow 0$ when $\sigma \rightarrow 0$. Using Corollary 4.9, we deduce easily that $J_\sigma(\mathbf{u}_\sigma) \rightarrow J(\tilde{\mathbf{u}})$. On the other hand, it is immediate that $\mathbf{u}_\sigma \in \mathcal{U}_{\sigma, ad}$ for every σ . Then the optimality of $\bar{\mathbf{u}}_\sigma$ and the definition of J_σ lead to

$$\frac{\lambda}{2} \|\bar{\mathbf{u}}_\sigma\|^2 \leq J_\sigma(\bar{\mathbf{u}}_\sigma) \leq J_\sigma(\mathbf{u}_\sigma) \leq C \quad \forall \sigma.$$

Therefore, we deduce the existence of subsequences weakly convergent. Let $\bar{\mathbf{u}}$ be one of these limit points. Obviously the property $\bar{\mathbf{u}} \in \mathcal{U}_{ad}$ holds. Moreover, using again Corollary 4.9 and the convexity of the cost functional in the third term involving the control, we have

$$\inf(P) \leq J(\bar{\mathbf{u}}) \leq \liminf_{\sigma \rightarrow 0} J_\sigma(\bar{\mathbf{u}}_\sigma) \leq \limsup_{\sigma \rightarrow 0} J_\sigma(\bar{\mathbf{u}}_\sigma) \leq \limsup_{\sigma \rightarrow 0} J_\sigma(\mathbf{u}_\sigma) = J(\tilde{\mathbf{u}}) = \inf(P),$$

which implies that $\bar{\mathbf{u}}$ is a solution of (P) as well as the convergence $J_\sigma(\bar{\mathbf{u}}_\sigma) \rightarrow J(\bar{\mathbf{u}})$. From this convergence along with the convergence properties of $\mathbf{y}_{\bar{\mathbf{u}}_\sigma} \rightarrow \mathbf{y}_{\bar{\mathbf{u}}}$ given in Corollary 4.9, we get $\|\bar{\mathbf{u}}_\sigma\|_{L^2(0, T; \mathbf{L}^2(\Omega_h))} \rightarrow \|\bar{\mathbf{u}}\|_{L^2(0, T; \mathbf{L}^2(\Omega))}$. Invoking once again (4.1), we obtain the strong convergence of $\{\bar{\mathbf{u}}_\sigma\}_\sigma$ to $\bar{\mathbf{u}}$ stated in (4.32). \square

The next theorem is important from a practical point of view because it states that every strict local minimum of problem (P) can be approximated by local minima of problems (P_σ) .

THEOREM 4.15. *Let $\bar{\mathbf{u}}$ be a strict local minimum of (P). Then there exists a sequence $\{\bar{\mathbf{u}}_\sigma\}_\sigma$ of local minima of problems (P_σ) such that (4.32) holds.*

Proof. Let $\bar{\mathbf{u}}$ be a strict local minimum of (P). Then there exists $\varepsilon > 0$ such that $\bar{\mathbf{u}}$ is the unique solution of

$$(P_\varepsilon) \quad \min_{\mathbf{u} \in \mathcal{U}_{ad} \cap \bar{B}_\varepsilon(\bar{\mathbf{u}})} J(\mathbf{u}).$$

Let us extend all the elements of \mathcal{U}_σ to $(0, T) \times \Omega$ by taking $\mathbf{u}_\sigma(t, x) = \bar{\mathbf{u}}(t, x)$ for any $(t, x) \in (0, T) \times (\Omega \setminus \Omega_h)$. Now, we consider the discrete problems

$$(P_{\varepsilon, \sigma}) \quad \min_{\mathbf{u}_\sigma \in \mathcal{U}_{\sigma, ad} \cap \bar{B}_\varepsilon(\bar{\mathbf{u}})} J_\sigma(\mathbf{u}_\sigma).$$

For every σ sufficiently small, the problem $(P_{\varepsilon, \sigma})$ has at least one solution. Indeed, the only delicate point is to check that $\mathcal{U}_{\sigma, ad} \cap \bar{B}_\varepsilon(\bar{\mathbf{u}})$ is not empty. To this end, we define $\mathbf{u}_\sigma \in \mathcal{U}_{\sigma, ad}$ as in (4.33), with $\tilde{\mathbf{u}}$ replaced by $\bar{\mathbf{u}}$. Then $\|\bar{\mathbf{u}} - \mathbf{u}_\sigma\|_{L^2(0, T; \mathbf{L}^2(\Omega))} \rightarrow 0$, and therefore $\mathbf{u}_\sigma \in \mathcal{U}_{\sigma, ad} \cap \bar{B}_\varepsilon(\bar{\mathbf{u}})$ for any σ sufficiently small. Let $\bar{\mathbf{u}}_\sigma$ be a solution of (P_ε) . Then we can argue as in the proof of Theorem 4.13 to deduce that any subsequence of $\{\bar{\mathbf{u}}_\sigma\}_\sigma$ converges strongly in $L^2(0, T; \mathbf{L}^2(\Omega))$ to a solution of (P_ε) . Since this problem has a unique solution, we have $\|\bar{\mathbf{u}} - \bar{\mathbf{u}}_\sigma\|_{L^2(0, T; \mathbf{L}^2(\Omega))} \rightarrow 0$ for the whole sequence as $\sigma \rightarrow 0$. This implies that the constraint $\bar{\mathbf{u}}_\sigma \in \bar{B}_\varepsilon(\bar{\mathbf{u}})$ is not active for σ small, and hence $\bar{\mathbf{u}}_\sigma$ is a local solution of (P_σ) and (4.32) is fulfilled. \square

4.4. Error estimates. We still assume that (4.21) holds. In this section $\bar{\mathbf{u}}$ will denote a local solution of problem (P), and for every σ , $\bar{\mathbf{u}}_\sigma$ denotes a local solution of (P_σ) such that $\|\bar{\mathbf{u}} - \bar{\mathbf{u}}_\sigma\|_{L^2(0, T; \mathbf{L}^2(\Omega_h))} \rightarrow 0$; see Theorems 4.13 and 4.15. Hereafter, all the elements $\mathbf{u} \in \mathcal{U}_\sigma$ are extended to $(0, T) \times \Omega$ by setting $\mathbf{u}(t, x) = \bar{\mathbf{u}}(t, x)$ for $(t, x) \in (0, T) \times (\Omega \setminus \Omega_h)$. We will also denote by $\bar{\mathbf{y}}$ and $\bar{\varphi}$ the state and adjoint state associated to $\bar{\mathbf{u}}$, and $\bar{\mathbf{y}}_\sigma$ and $\bar{\varphi}_\sigma$ will denote the discrete state and adjoint state associated to $\bar{\mathbf{u}}_\sigma$. The goal of this section is to prove the following theorem.

THEOREM 4.16. *Suppose that (3.18) holds. Then there exists a constant $C > 0$ independent of σ such that*

$$(4.34) \quad \|\bar{\mathbf{u}} - \bar{\mathbf{u}}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))} \leq Ch,$$

$$(4.35) \quad \|\bar{\mathbf{y}} - \bar{\mathbf{y}}_\sigma\|_{L^\infty(0,T;\mathbf{L}^2(\Omega_h))} + \|\bar{\mathbf{y}} - \bar{\mathbf{y}}_\sigma\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))} \leq Ch,$$

$$(4.36) \quad \|\bar{\boldsymbol{\varphi}} - \bar{\boldsymbol{\varphi}}_\sigma\|_{L^\infty(0,T;\mathbf{L}^2(\Omega_h))} + \|\bar{\boldsymbol{\varphi}} - \bar{\boldsymbol{\varphi}}_\sigma\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))} \leq Ch.$$

The estimates (4.35) and (4.36) are an immediate consequence of (4.34), (4.22), and (4.31). We only have to prove (4.34). To this end, we proceed by contradiction and assume that it is false. This implies that

$$\limsup_{\sigma \rightarrow 0} \frac{1}{h} \|\bar{\mathbf{u}} - \bar{\mathbf{u}}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))} = +\infty;$$

therefore, there exists a sequence of σ such that

$$(4.37) \quad \lim_{\sigma \rightarrow 0} \frac{1}{h} \|\bar{\mathbf{u}} - \bar{\mathbf{u}}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))} = +\infty.$$

We will obtain a contradiction for this sequence. We need some lemmas. The first one is concerned with the projection of $\bar{\mathbf{u}}$ on \mathcal{U}_σ given by the formulas (4.33) and denoted in what follows by \mathbf{u}_σ . Let us recall that according to Theorem 3.3 the regularity $\bar{\mathbf{u}} \in \mathbf{H}^1(\Omega_T)$ holds for any local minimum.

LEMMA 4.17. *There exists a constant $C > 0$ independent of σ such that*

$$(4.38) \quad \|\bar{\mathbf{u}} - \mathbf{u}_\sigma\|_{\mathbf{H}^1(\Omega_{Th})^*} + h \|\bar{\mathbf{u}} - \mathbf{u}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))} \leq Ch^2 \|\bar{\mathbf{u}}\|_{\mathbf{H}^1(\Omega_T)},$$

where $\Omega_{Th} = (0, T) \times \Omega_h$.

Proof. The estimate in the $L^2(0, T; \mathbf{L}^2(\Omega_h))$ -norm is well known. Let us check the estimate in the $\mathbf{H}^1(\Omega_{Th})^*$ -norm. Let $\mathbf{v} \in \mathbf{H}^1(\Omega_{Th})$ be any element and take \mathbf{v}_σ as the projection according to the expression (4.33). From the definition of the projection we have

$$\begin{aligned} \int_0^T \int_{\Omega_h} \mathbf{v}(\bar{\mathbf{u}} - \mathbf{u}_\sigma) \, dt dx &= \int_0^T \int_{\Omega_h} (\mathbf{v} - \mathbf{v}_\sigma)(\bar{\mathbf{u}} - \mathbf{u}_\sigma) \, dt dx \\ &\leq \|\mathbf{v} - \mathbf{v}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))} \|\bar{\mathbf{u}} - \mathbf{u}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))} \leq Ch^2 \|\bar{\mathbf{u}}\|_{\mathbf{H}^1(\Omega_T)} \|\mathbf{v}\|_{\mathbf{H}^1(\Omega_{Th})}, \end{aligned}$$

which proves the lemma. \square

Since $\bar{\mathbf{u}}_\sigma$ is a local minimum of (P_σ) , J_σ is a C^∞ mapping and $\mathbf{u}_\sigma \in \mathcal{U}_{\sigma,ad}$. Then $J'_\sigma(\bar{\mathbf{u}}_\sigma)(\mathbf{u}_\sigma - \bar{\mathbf{u}}_\sigma) \geq 0$. This inequality can be rewritten in the form

$$\begin{aligned} J'(\bar{\mathbf{u}}_\sigma)(\bar{\mathbf{u}} - \bar{\mathbf{u}}_\sigma) &+ [J'_\sigma(\bar{\mathbf{u}}_\sigma) - J'(\bar{\mathbf{u}}_\sigma)](\bar{\mathbf{u}} - \bar{\mathbf{u}}_\sigma) \\ &+ [J'_\sigma(\bar{\mathbf{u}}_\sigma) - J'(\bar{\mathbf{u}})](\mathbf{u}_\sigma - \bar{\mathbf{u}}) + J'(\bar{\mathbf{u}})(\mathbf{u}_\sigma - \bar{\mathbf{u}}) \geq 0. \end{aligned}$$

On the other hand, since $\bar{\mathbf{u}}_\sigma \in \mathcal{U}_{ad}$, then $J'(\bar{\mathbf{u}})(\bar{\mathbf{u}}_\sigma - \bar{\mathbf{u}}) \geq 0$. Adding this inequality to the last one, we obtain

$$(4.39) \quad \begin{aligned} [J'(\bar{\mathbf{u}}_\sigma) - J'(\bar{\mathbf{u}})](\bar{\mathbf{u}}_\sigma - \bar{\mathbf{u}}) &\leq [J'_\sigma(\bar{\mathbf{u}}_\sigma) - J'(\bar{\mathbf{u}}_\sigma)](\bar{\mathbf{u}} - \bar{\mathbf{u}}_\sigma) \\ &+ [J'_\sigma(\bar{\mathbf{u}}_\sigma) - J'(\bar{\mathbf{u}})](\mathbf{u}_\sigma - \bar{\mathbf{u}}) + J'(\bar{\mathbf{u}})(\mathbf{u}_\sigma - \bar{\mathbf{u}}). \end{aligned}$$

This inequality is crucial in the proof. First, we get an estimate from below for the left-hand side, then we estimate from above the three terms on the right-hand side.

LEMMA 4.18. Suppose that $\{\bar{\mathbf{u}}_\sigma\}_\sigma$ satisfies (4.37) and let $\delta > 0$ be given by Remark 3.5. Then there exists σ_0 such that

$$(4.40) \quad \frac{1}{2} \min\{\delta, \lambda\} \|\bar{\mathbf{u}}_\sigma - \bar{\mathbf{u}}\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))}^2 \leq [J'(\bar{\mathbf{u}}_\sigma) - J'(\bar{\mathbf{u}})](\bar{\mathbf{u}}_\sigma - \bar{\mathbf{u}}) \quad \text{if } |\sigma| < |\sigma_0|.$$

Proof. In this proof, we follow the steps of [7, Lemma 7.2]. Applying the mean value theorem we get for some $\hat{\mathbf{u}}_\sigma = \bar{\mathbf{u}} + \theta_h(\bar{\mathbf{u}}_\sigma - \bar{\mathbf{u}})$ that

$$(4.41) \quad [J'(\bar{\mathbf{u}}_\sigma) - J'(\bar{\mathbf{u}})](\bar{\mathbf{u}}_\sigma - \bar{\mathbf{u}}) = J''(\hat{\mathbf{u}}_\sigma)(\bar{\mathbf{u}}_\sigma - \bar{\mathbf{u}})^2.$$

Set $\rho_\sigma = \|\bar{\mathbf{u}}_\sigma - \bar{\mathbf{u}}\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))}$ and $\mathbf{v}_\sigma = \frac{1}{\rho_\sigma}(\bar{\mathbf{u}}_\sigma - \bar{\mathbf{u}})$. Taking a subsequence, we can assume that $\mathbf{v}_\sigma \rightharpoonup \mathbf{v}$ in $L^2(0,T;\mathbf{L}^2(\Omega_h))$. Let us prove that \mathbf{v} belongs to the critical cone $\mathcal{C}_{\bar{\mathbf{u}}}$ defined in (3.11). First, remark that every \mathbf{v}_σ satisfies the sign conditions (3.12)–(3.13), and hence \mathbf{v} also does. Now, we show that $v_j(t, x) = 0$ if $d_j(t, x) \neq 0$, $\bar{\mathbf{d}}$ being defined by (3.10). Denote $\bar{\mathbf{d}}_\sigma = J'_\sigma(\bar{\mathbf{u}}_\sigma) = \bar{\varphi}_\sigma + \lambda \bar{\mathbf{u}}_\sigma$; see (4.26). Observe that $\|\bar{\mathbf{u}} - \bar{\mathbf{u}}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))} \rightarrow 0$ and (4.31) imply the convergence $\|\bar{\mathbf{d}} - \bar{\mathbf{d}}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))} \rightarrow 0$. Now, we have

$$\begin{aligned} \int_0^T \int_\Omega \bar{\mathbf{d}} \mathbf{v} \, dx dt &= \lim_{\sigma \rightarrow 0} \int_0^T \int_{\Omega_h} \bar{\mathbf{d}}_\sigma \mathbf{v}_\sigma \, dx dt \\ &= \lim_{\sigma \rightarrow 0} \frac{1}{\rho_\sigma} \left\{ \int_0^T \int_{\Omega_h} \bar{\mathbf{d}}_\sigma (\mathbf{u}_\sigma - \bar{\mathbf{u}}) \, dx dt + \int_0^T \int_{\Omega_h} \bar{\mathbf{d}}_\sigma (\bar{\mathbf{u}}_\sigma - \mathbf{u}_\sigma) \, dx dt \right\}. \end{aligned}$$

From (4.37), (4.38), and the inequality $J'_\sigma(\bar{\mathbf{u}}_\sigma)(\mathbf{u}_\sigma - \bar{\mathbf{u}}_\sigma) \geq 0$ we conclude that

$$\begin{aligned} \int_0^T \int_\Omega \bar{\mathbf{d}} \mathbf{v} \, dx dt &\leq \lim_{\sigma \rightarrow 0} \frac{1}{\rho_\sigma} \int_0^T \int_{\Omega_h} \bar{\mathbf{d}}_\sigma (\mathbf{u}_\sigma - \bar{\mathbf{u}}) \, dx dt \\ &\leq \lim_{\sigma \rightarrow 0} \frac{Ch}{\|\bar{\mathbf{u}}_\sigma - \bar{\mathbf{u}}\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))}} = 0. \end{aligned}$$

Since \mathbf{v} satisfies the sign conditions (3.12)–(3.13), then $d_j(t, x)v_j(t, x) \geq 0$; hence the above inequality implies that (3.14) holds as well, and then $\mathbf{v} \in \mathcal{C}_{\bar{\mathbf{u}}}$. Now, from the definition of \mathbf{v}_σ , (3.4), and (3.20) we get

$$\begin{aligned} &\lim_{\sigma \rightarrow 0} J''(\hat{\mathbf{u}}_\sigma)(\bar{\mathbf{u}}_\sigma - \bar{\mathbf{u}})^2 \\ &= \lim_{\sigma \rightarrow 0} \left\{ \int_0^T \int_\Omega (|\mathbf{z}_{\mathbf{v}_\sigma}|^2 - 2(\mathbf{z}_{\mathbf{v}_\sigma} \cdot \nabla) \mathbf{z}_{\mathbf{v}_\sigma} \varphi_{\hat{\mathbf{u}}_\sigma}) \, dx dt + \gamma \int_\Omega |\mathbf{z}_{\mathbf{v}_\sigma}(T)|^2 \, dx + \lambda \right\} \\ &= \int_0^T \int_\Omega (|\mathbf{z}_{\mathbf{v}}|^2 - 2(\mathbf{z}_{\mathbf{v}} \cdot \nabla) \mathbf{z}_{\mathbf{v}} \bar{\varphi}) \, dx dt + \gamma \int_\Omega |\mathbf{z}_{\mathbf{v}}(T)|^2 \, dx + \lambda \\ &= J''(\bar{\mathbf{u}}) \mathbf{v}^2 + \lambda \left(1 - \|\mathbf{v}\|_{L^2(0,T;\mathbf{L}^2(\Omega))}^2 \right) \geq \lambda + (\delta - \lambda) \|\mathbf{v}\|_{L^2(0,T;\mathbf{L}^2(\Omega))}^2. \end{aligned}$$

Taking into account that $\|\mathbf{v}\|_{L^2(0,T;\mathbf{L}^2(\Omega))} \leq 1$, these inequalities lead to

$$\lim_{\sigma \rightarrow 0} J''(\hat{\mathbf{u}}_\sigma) \mathbf{v}_\sigma^2 \geq \min\{\delta, \lambda\} > 0,$$

which proves the existence of σ_0 such that $J''(\hat{\mathbf{u}}_\sigma) \mathbf{v}_\sigma^2 \geq \frac{1}{2} \min\{\delta, \lambda\} \forall |\sigma| < |\sigma_0|$. From this inequality, the definition of \mathbf{v}_σ , and (4.41) we deduce (4.40). \square

With (4.39) and (4.40) we obtain

$$(4.42) \quad \begin{aligned} & \frac{1}{2} \min\{\delta, \lambda\} \|\bar{\mathbf{u}}_\sigma - \bar{\mathbf{u}}\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))}^2 \leq [J'_\sigma(\bar{\mathbf{u}}_\sigma) - J'(\bar{\mathbf{u}}_\sigma)](\bar{\mathbf{u}} - \bar{\mathbf{u}}_\sigma) \\ & + [J'_\sigma(\bar{\mathbf{u}}_\sigma) - J'(\bar{\mathbf{u}})](\mathbf{u}_\sigma - \bar{\mathbf{u}}_\sigma) + J'(\bar{\mathbf{u}})(\mathbf{u}_\sigma - \bar{\mathbf{u}}). \end{aligned}$$

Let us estimate the three terms on the right-hand side. From (3.3) and (4.26), along with the fact that $\bar{\mathbf{u}} = \bar{\mathbf{u}}_\sigma$ in $(0, T) \times (\Omega \setminus \Omega_h)$, we have

$$|[J'_\sigma(\bar{\mathbf{u}}_\sigma) - J'(\bar{\mathbf{u}}_\sigma)](\bar{\mathbf{u}} - \bar{\mathbf{u}}_\sigma)| \leq \|\varphi_{\bar{\mathbf{u}}_\sigma} - \bar{\varphi}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))} \|\bar{\mathbf{u}} - \bar{\mathbf{u}}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))}.$$

Taking $\mathbf{u} = \mathbf{v} = \bar{\mathbf{u}}_\sigma$ in (4.31), the previous inequality leads to

$$(4.43) \quad |[J'_\sigma(\bar{\mathbf{u}}_\sigma) - J'(\bar{\mathbf{u}}_\sigma)](\bar{\mathbf{u}} - \bar{\mathbf{u}}_\sigma)| \leq Ch \|\bar{\mathbf{u}} - \bar{\mathbf{u}}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))}.$$

For the second term of (4.42) we use again (4.31), with $\mathbf{u} = \bar{\mathbf{u}}$ and $\mathbf{v} = \bar{\mathbf{u}}_\sigma$, and (4.38) to obtain

$$(4.44) \quad \begin{aligned} & |[J'_\sigma(\bar{\mathbf{u}}_\sigma) - J'(\bar{\mathbf{u}})](\mathbf{u}_\sigma - \bar{\mathbf{u}}_\sigma)| \\ & \leq \{\|\bar{\varphi}_\sigma - \bar{\varphi}\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))} + \|\bar{\mathbf{u}}_\sigma - \bar{\mathbf{u}}\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))}\} \|\mathbf{u}_\sigma - \bar{\mathbf{u}}\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))} \\ & \leq C \{h + \|\bar{\mathbf{u}} - \bar{\mathbf{u}}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))}\} h. \end{aligned}$$

Last, we estimate the third term using again (4.38):

$$(4.45) \quad |J'(\bar{\mathbf{u}})(\mathbf{u}_\sigma - \bar{\mathbf{u}})| \leq \|\bar{\varphi} + \lambda \bar{\mathbf{u}}\|_{\mathbf{H}^1(\Omega_{Th})} \|\mathbf{u}_\sigma - \bar{\mathbf{u}}\|_{\mathbf{H}^1(\Omega_{Th})^*} \leq Ch^2.$$

Finally (4.34) follows from (4.42)–(4.45) with the help of Young's inequality, which contradicts (4.37).

REFERENCES

- [1] F. ABERGEL AND R. TEMAM, *On some control problems in fluid mechanics*, Theoret. Comput. Fluid Dynam., 1 (1990), pp. 303–325.
- [2] J. F. BONNANS AND H. ZIDANI, *Optimal control problems with partially polyhedral constraints*, SIAM J. Control Optim., 37 (1999), pp. 1726–1741.
- [3] S. C. BRENNER AND L. R. SCOTT, *The Mathematical Theory of Finite Element Methods*, Springer-Verlag, New York, Berlin, Heidelberg, 1994.
- [4] E. CASAS, *An optimal control problem governed by the evolution Navier–Stokes equations*, in Optimal Control of Viscous Flows, S. S. Sritharan, ed., SIAM, Philadelphia, 1998, pp. 79–95.
- [5] E. CASAS AND M. MATEOS, *Second order optimality conditions for semilinear elliptic control problems with finitely many state constraints*, SIAM J. Control Optim., 40 (2002), pp. 1431–1454.
- [6] E. CASAS, M. MATEOS, AND J.-P. RAYMOND, *Error estimates for the numerical approximation of a distributed control problem for the steady-state Navier–Stokes equations*, SIAM J. Control Optim., 46 (2007), pp. 952–982.
- [7] E. CASAS AND J.-P. RAYMOND, *Error estimates for the numerical approximation of Dirichlet boundary control for semilinear elliptic equations*, SIAM J. Control Optim., 45 (2006), pp. 1586–1611.
- [8] K. CHRYSAFINOS, *Convergence of discontinuous Galerkin approximations of an optimal control problem associated to semilinear parabolic PDE's*, M2AN Math. Model. Numer. Anal., 44 (2010), pp. 189–206.
- [9] K. CHRYSAFINOS AND N. J. WALKINGTON, *Discontinuous Galerkin approximations of the Stokes and Navier–Stokes equations*, Math. Comp., 79 (2010), pp. 2135–2167.
- [10] K. DECKELNICK AND M. HINZE, *Semidiscretization and error estimates for distributed control of the instationary Navier–Stokes equations*, Numer. Math., 97 (2004), pp. 297–320.

- [11] K. ERIKSSON AND C. JOHNSON, *Adaptive finite element methods for parabolic problems I: A linear model problem*, SIAM J. Numer. Anal., 28 (1991), pp. 43–77.
- [12] K. ERIKSSON AND C. JOHNSON, *Adaptive finite element methods for parabolic problems II: Optimal error estimates in $L_\infty(L^2)$ and $L_\infty(L_\infty)$* , SIAM J. Numer. Anal., 32 (1995), pp. 706–740.
- [13] K. ERIKSSON AND C. JOHNSON, *Adaptive finite element methods for parabolic problems IV: Non-linear problems*, SIAM J. Numer. Anal., 32 (1995), pp. 1729–1749.
- [14] D. ESTEP AND S. LARSSON, *The discontinuous Galerkin method for semilinear parabolic equations*, RAIRO Modél. Math. Anal. Numér., 27 (1993), pp. 35–54.
- [15] P. GIRAULT AND P. A. RAVIART, *Finite Element Methods for Navier-Stokes Equations. Theory and Algorithms*, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1986.
- [16] M. D. GUNZBURGER, *Perspectives in Flow Control and Optimization*, Adv. Des. Control 5, SIAM, Philadelphia, 2003.
- [17] M. D. GUNZBURGER, *Flow Control*, Springer-Verlag, New York, 1995.
- [18] M. D. GUNZBURGER AND S. MANSERVISI, *The velocity tracking problem for Navier–Stokes flows with bounded distributed control*, SIAM J. Control Optim., 37 (1999), pp. 1913–1945.
- [19] M. D. GUNZBURGER AND S. MANSERVISI, *Analysis and approximation of the velocity tracking problem for Navier–Stokes flows with distributed control*, SIAM J. Numer. Anal., 37 (2000), pp. 1481–1512.
- [20] M. HINZE AND K. KUNISCH, *Second order methods for optimal control of time-dependent fluid flow*, SIAM J. Control Optim., 40 (2001), pp. 925–946.
- [21] O. A. LADYZHENSKAYA, *The Mathematical Theory of Viscous Incompressible Flow*, 2nd ed., Gordon and Breach, New York, 1969.
- [22] J. L. LIONS, *Quelques Méthodes de Résolution des Problèmes aux Limites non Linéaires*, Dunod, Paris, 1969.
- [23] J. L. LIONS AND E. MAGENES, *Problèmes aux Limites non Homogènes*, Dunod, Paris, 1968.
- [24] D. MEIDNER AND B. VEXLER, *A priori error estimates for the space-time finite element discretization of parabolic optimal control problems. Part I: Problems without control constraints*, SIAM J. Control Optim., 47 (2008), pp. 1150–1177.
- [25] D. MEIDNER AND B. VEXLER, *A priori error estimates for the space-time finite element discretization of parabolic optimal control problems. Part II: Problems with control constraints*, SIAM J. Control Optim., 47 (2008), pp. 1301–1329.
- [26] D. MEIDNER AND B. VEXLER, *A priori error analysis of the Petrov–Galerkin Crank–Nicolson scheme for parabolic optimal control problems*, SIAM J. Control Optim., 49 (2011), pp. 2183–2211.
- [27] I. NEITZEL AND B. VEXLER, *A priori error estimates for space-time finite element discretization of semilinear parabolic optimal control problems*, Numer. Math., 120 (2012), pp. 345–386.
- [28] P. A. RAVIART AND J. M. THOMAS, *Introduction à L’analyse Numérique des Equations aux Dérivées Partielles*, Masson, Paris, 1983.
- [29] A. RÖSCH, *Error estimates for parabolic optimal control problems with control constraints*, Z. Anal. Anwendungen, 23 (2004), pp. 353–376.
- [30] S. S. SRITHARAN, ED., *Optimal Control of Viscous Flow*, SIAM, Philadelphia, 1998.
- [31] R. TEMAM, *Navier-Stokes Equations*, North-Holland, Amsterdam, 1979.
- [32] V. THOMÉE, *Galerkin Finite Element Methods for Parabolic Problems*, Springer-Verlag, Berlin, 1997.
- [33] F. TRÖLTZSCH, *Optimal Control of Partial Differential Equations*, Grad. Stud. Math. 112, AMS, Providence, RI, 2010.
- [34] F. TRÖLTZSCH AND D. WACHSMUTH, *Second-order sufficient optimality conditions for the optimal control of Navier-Stokes equations*, ESAIM Control Optim. Calc. Var., 12 (2006), pp. 93–119.